Lectures on
Semi-group Theory and its
Application to Cauchy’s Problem
in Partial Differential Equations

By
K. Yosida

Tata Institute of Fundamental Research, Bombay
1957
Lectures on
Semi-group Theory and its
Application to Cauchy’s Problem
in Partial Differential Equations

By
K. Yosida

Notes by
M.S. Narasimhan

Tata Institute of Fundamental Research,
Bombay
1957
Contents

1 Lecture 1 1
  1 Introduction ............................................. 1

I Survey of some basic concepts and ... 3
  2 Normed linear spaces: ................................. 5
  3 Pre-Hilbert spaces .................................... 6
  4 Example of a pre-Hilbert space ...................... 7
  5 Banach spaces ........................................... 8
  6 Hilbert space ........................................... 8
  7 Example of Banach spaces ............................. 8
  8 Example of a Hilbert space ............................ 9
  9 Completion of a normed linear space ................. 10
 10 Additive operators .................................... 10

2 Lecture 2 13
  1 Linear operators ....................................... 13
  2 Hahn-Banach lemma ................................... 13
  3 Lemma (Hahn-Banach) .................................. 14
  4 Hahn-Banach extension theorem ...................... 15
  5 Hahn-Banach extension theorem for ................ 15
  6 Existence of non-trivial linear functionals .......... 17
  7 Orthogonal projection and ............................ 18
## Contents

### 3 Lecture 3
1. The Conjugate space (dual) of a normed linear space . . . 21
2. The Resonance Theorem . . . . . . . . . . . . . . . . . 21
3. Weak convergence . . . . . . . . . . . . . . . . . . . 23
4. A counter-example . . . . . . . . . . . . . . . . . . . 24

### 4 Lecture 4
1. Local weak compactness of a Hilbert space . . . . . . . 27
2. Lax-Milgram theorem . . . . . . . . . . . . . . . . . . 28

### II Semi-group Theory

### 5 Lecture 5
1. Some examples of semi-groups . . . . . . . . . . . . . 35
2. The infinitesimal generator of a semi-group . . . . . . . 39

### 6 Lecture 6
1. The resolvent set and... . . . . . . . . . . . . . . . . . . 46
2. Examples . . . . . . . . . . . . . . . . . . . . . . . . 47

### 7 Lecture 7
1. The exponential of a linear operator . . . . . . . . . . . 53
2. Representation of semi-groups . . . . . . . . . . . . . 56

### 8 Lecture 8
1. An application of the representation theorem . . . . . . . 59
2. Characterization of the... . . . . . . . . . . . . . . . . . 59

### 9 Lecture 9
1. Group of operators . . . . . . . . . . . . . . . . . . . 65

### 10 Lecture 10
1. Supplementary results . . . . . . . . . . . . . . . . . . . 69
## Contents

<table>
<thead>
<tr>
<th>Lecture</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>11</strong> Lecture 11</td>
<td>75</td>
</tr>
<tr>
<td>1 Temporally homogeneous Markov...</td>
<td>75</td>
</tr>
<tr>
<td>2 Brownian motion on a homogeneous Riemannian space</td>
<td>76</td>
</tr>
<tr>
<td><strong>12</strong> Lecture 12</td>
<td>81</td>
</tr>
<tr>
<td>1 Brownian motion on a homogeneous...</td>
<td>81</td>
</tr>
<tr>
<td><strong>III</strong> Regularity properties of solutions of linear elliptic...</td>
<td>85</td>
</tr>
<tr>
<td><strong>13</strong> Lecture 13</td>
<td>87</td>
</tr>
<tr>
<td>1 Strong differentiability</td>
<td>87</td>
</tr>
<tr>
<td>2 Weak solutions of linear differential operators</td>
<td>88</td>
</tr>
<tr>
<td>3 Elliptic operators</td>
<td>89</td>
</tr>
<tr>
<td>4 Fourier Transforms</td>
<td>89</td>
</tr>
<tr>
<td><strong>14</strong> Lecture 14</td>
<td>93</td>
</tr>
<tr>
<td>1 Garding’s inequality</td>
<td>93</td>
</tr>
<tr>
<td><strong>15</strong> Lecture 15</td>
<td>99</td>
</tr>
<tr>
<td>1 Proof of the Friedrichs - Lax - Nirenberg theorem</td>
<td>99</td>
</tr>
<tr>
<td><strong>16</strong> Lecture</td>
<td>103</td>
</tr>
<tr>
<td>1 Proof of Lemma 3</td>
<td>103</td>
</tr>
<tr>
<td><strong>17</strong> Lecture 17</td>
<td>107</td>
</tr>
<tr>
<td>1 Proof of Lemma 2</td>
<td>107</td>
</tr>
<tr>
<td>2 Proof of Lemma 1</td>
<td>108</td>
</tr>
<tr>
<td><strong>IV</strong> Application of the semi-group theory...</td>
<td>113</td>
</tr>
<tr>
<td><strong>18</strong> Lecture 18</td>
<td>115</td>
</tr>
<tr>
<td>1 Cauchy problem for the diffusion equation</td>
<td>115</td>
</tr>
<tr>
<td>2 Garding’s inequality</td>
<td>118</td>
</tr>
</tbody>
</table>
19 Lecture 19 121
   1 The Cauchy problem for the wave equation . . . . . . . 121

20 Lecture 20 127
   1 Cauchy problem for the wave equation (continued) . . . 127

21 Lecture 21 133
   1 Integration of the Fokker-Planck equation . . . . . . . 133

22 Lecture 22 139
   1 Integration of the Fokker-Planck... . . . . . . . . . . . . 139

23 Lecture 23 145
   1 Integration of the Fokker-Planck equation (Contd.) ... . 145
Lecture 1

1 Introduction

The analytical theory of one-parameter semi-groups deals with the exponential function in infinite dimensional function spaces. It is a natural generalization of the theorem of Stone on one-parameter groups of unitary operators in a Hilbert space.

In these lectures, we shall be concerned with the differentiability and the representation of one-parameter semi-groups of bounded linear operators on a Banach space and with some of their applications to the initial value problem (Cauchy’s problem) for differential equations, especially for the diffusion equation (heat equation) and the wave equation.

The ordinary exponential function solves the initial value problem:

\[ \frac{dy}{dt} = \alpha y, \quad y(0) = C. \]

We consider the diffusion equation

\[ \frac{\partial u}{\partial t} = \Delta u, \]

where \( \Delta = \sum_{i=1}^{m} \frac{\partial^2}{\partial x_i^2} \) is the Laplacian in the Euclidean m-space \( E^m \); we wish to find a solution \( u = u(x,t), t \geq 0 \), of this equation satisfying the initial condition \( u(x,0) = f(x) \), where \( f(x) = f(x_1, \ldots, x_n) \) is a given
function of $x$. We shall also study the wave equation

\[ \frac{\partial^2 u}{\partial t^2} = \Delta u, \quad -\infty \leq t \leq \infty \]

with the initial data

\[ u(x, 0) = f(x) \text{ and } \left( \frac{\partial u}{\partial t} \right)_{t=0} = g(x), \]

$f$ and $g$ being given functions. This may be written in the vector form as follows:

\[ \frac{\partial}{\partial t} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \quad v = \frac{\partial u}{\partial t} \]

with the initial condition

\[ \begin{pmatrix} u(0) \\ v(0) \end{pmatrix} = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}. \]

So in a suitable function space the wave equation is of the same form as the heat equation - differentiation with respect to the time parameter on the left and another operator on the right - or again similar to the equation \( \frac{dy}{dt} = \alpha y \). Since the solution in the last case is the exponential function, it is suggested that the heat equation and the wave equation may be solved by properly defining the exponential functions of the operators $\Delta$ and $\begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}$ in suitable function spaces. This is the motivation for the application of the semi-group theory to Cauchy’s problem.

Our method will give an explanation why in the case of the heat equation the time parameter is restricted to non-negative values, while in the case of the wave equation it may extend between $-\infty$ and $\infty$.

Before going into the details, we give a survey of some of the basic concepts and results from the theory of Banach spaces and Hilbert spaces.
Part I

Survey of some basic concepts and results from the theory of Banach spaces
2. Normed linear spaces:

**Definition.** A set $X$ is called a linear space over a field $K$ if the following conditions are satisfied:

1) $X$ is an abelian group (written additively).

2) There is defined a scalar multiplication: to every element $x$ of $X$ and each $\alpha \in K$ there is associated an element of $X$, denoted by $\alpha x$, such that

$$(\alpha + \beta)x = \alpha x + \beta x, \quad \alpha, \beta \in K, \quad x \in X$$

$$\alpha(x + y) = \alpha x + \alpha y, \quad \alpha \in K, x, y \in X$$

$$(\alpha \beta)x = \alpha(\beta x)$$

$1x = x, \quad 1 \in K$$is the unit element of $K$.

We shall denote by Greek letters the elements of $K$ and by Roman letters the elements of $X$. The zero of $X$ and the zero of $K$ will both be denoted by 0. We have $0x = 0$.

In the sequel we consider linear spaces only over the real number field or the complex number field. A linear space will be said to be real or complex according as the field is the real number field or the complex number field. In what follows, by a linear space we always mean a real or a complex linear space.

**Definition.** A subset $M$ of a linear space $X$ is called a linear subspace (or a subspace) if whenever $x, y \in M$ and $\alpha, \beta \in K$ then $\alpha x + \beta y \in M$.

2 Normed linear spaces:

**Definition.** A linear space $X$ (real or complex) is called a normed linear space if, for every $x \in X$ there is associated a real number, denoted by $||x||$, such that

i) $||x|| \geq 0$ and $||x|| = 0$ if and only if $x = 0$.

ii) $||\alpha x|| = |\alpha||x||$, ($\alpha$ is a scalar and $|\alpha|$ is the modulus of $\alpha$).

iii) $||x + y|| \leq ||x|| + ||y||$, $x, y \in X$ (triangle inequality). $||x||$ is called the norm of $x$. 
A normed linear space becomes a metric space if the distance $d(x, y)$ between two elements $x$ and $y$ is defined by $d(x, y) = ||x - y||$. We say that a sequence of elements $\{x_n\}$ of $X$ converges strongly to $x \in X$, and write $s - \lim_{n \to \infty} x_n = x$ (or simply $\lim x_n = x$), if $\lim ||x_n - x|| = 0$. (This limit, if it exists, is unique by the triangle inequality).

**Proposition.** If $\lim_{n \to \infty} \alpha_n = \alpha (\alpha_n, \alpha \in K)$, $s - \lim_{n \to \infty} x_n = x$ and $s - \lim_{n \to \infty} y_n = y$, then $s - \lim_{n \to \infty} \alpha_n x_n = \alpha x$ and $s - \lim_{n \to \infty} (x_n + y_n) = x + y$.

**Proof.**

$||(x_n + y_n) - (x + y)|| = ||(x_n - x) + (y_n - y)||$

$\leq ||(x_n - x)|| + ||(y_n - y)||$ (Triangle inequality)

$\to 0.$

$||\alpha_n x_n - \alpha x|| \leq ||\alpha x - \alpha_n x|| + ||\alpha_n x - \alpha_n x_n||$

$= |\alpha - \alpha_n|||x|| + |\alpha_n||x - x_n||$

$\to 0.$

**Proposition.** If $s - \lim_{n \to \infty} x_n = x$ then $\lim_{n \to \infty} ||x_n|| = ||x||$, i.e., norm is a continuous function.

**Proof.** We have, from the triangle inequality,

$||x|| - ||y|| \leq ||x - y||$

now take $y = x_n$ and let $n \to \infty$. \(\square\)

### 3 Pre-Hilbert spaces

A special class of normed linear spaces - pre-Hilbert spaces-will be of fundamental importance in our later discussion of differential equations. These normed linear spaces in which the norm is defined by scalar product.

**Definition.** A linear space $X$ is called a pre-Hilbert space if for every ordered pair of elements $(x, y) (x, y \in X)$ there is associated a number (real number if $X$ is a real linear space and complex number if $X$ is a complex linear space) such that
4. Example of a pre-Hilbert space

i) \((x, x) \geq 0\) and \((x, x) = 0\) if and only if \(x = 0\).

ii) \((\alpha x, y) = \alpha (x, y)\), for every number \(\alpha\).

iii) \((x, y) = (y, x)\) if \((y, x)\) denotes the complex conjugate of \((y, x)\).

iv) \((x + y, z) = (x, z) + (y, z)\) \(x, y, z \in X\).

\((x, y)\) is called the scalar product between \(x\) and \(y\).

If we define \(\|x\| = \sqrt{(x, x)}\), a pre-Hilbert space becomes a normed linear space, as is verified easily using Schwarz’s inequality proved below.

**Proposition.**

i) \(|(x, y)| \leq ||x|| ||y||\) (Schwarz’s inequality)

ii) \(||x + y||^2 + ||x - y||^2 = 2(||x||^2 + ||y||^2)\) (Euclidean property)

**Proof.** (ii) is easily verified. To prove (i), we observe that, for every real number \(\alpha\),

\[
0 \leq (x + \alpha(x, y)y, x + \alpha(x, y)y) = (x, x) + 2\alpha(x, y)^2 + \alpha^2(x, y)^2(y, y).
\]

This quadratic form in \(\alpha\), being always non-negative should have non-positive discriminant so that

\[
|(x, y)^2 - ||x||^2 ||y||^2 (x, y)^2| \leq 0.
\]

If \((x, y) = 0\), (i) is obviously satisfied; if \((x, y) \neq 0\), Schwarz’s inequality follows from the above inequality. \(\square\)

**4 Example of a pre-Hilbert space**

Let \(\mathbb{R}\) be a domain in Euclidean \(m\)-space \(E^m\). Let \(\mathcal{D}^k(\mathbb{R})\) denote the set of all complex valued functions \(f(x) = f(x_1, \ldots, x_n)\) which are of class \(C^k\) in \(\mathbb{R}\)(i.e., \(k\) times continuously differentiable) and which have compact support. These functions form a linear space with the ordinary function.
sum and scalar multiplication. Define the scalar product between two functions $f$ and $g$ by

$$(f, g)_k = \sum_{|n| \leq k} \int \mathcal{D}_{\mathcal{K}}(n) f(x) D^{(n)} g(x) dx, \quad 0 \leq k < \infty,$$

where $n = (n_1, \ldots, n_m)$ is a system of non-negative integers, $|n| = n_1 + + n_m$ and

$$D^{(n)} = \frac{\partial^{|n|}}{\partial x_1^{n_1} \partial x_2^{n_2} \cdots \partial x_m^{n_m}}.$$

## 5 Banach spaces

**Definition.** A normed linear space is called a Banach spaces if it is complete in the sense of the metric given by the norm.

(Completeness means that every Cauchy sequence is convergent: if \{x_n\} \subset X is any Cauchy sequence, i.e., a sequence \{x_n\} for which $\|x_m - x_n\| \to 0$ as $m, n \to \infty$ independently, then there exists an element $x \in X$ such that $\lim_{n \to \infty} \|x_n - x\| = 0$ is unique).

## 6 Hilbert space

**Definition.** A pre-Hilbert space which is complete (considered as a normed linear space) is called a Hilbert space.

The pre-Hilbert space $\mathcal{D}^\mathcal{K}(\mathbb{R})$ defined in the last example is not complete.

## 7 Example of Banach spaces

1) $C[\alpha, \beta]$: Let $[\alpha, \beta]$ be a closed interval $-\infty \leq \alpha < \beta \leq \infty$. Let $C[\alpha, \beta]$ denote the set of all bounded continuous complex-valued functions $x(t)$ on $[\alpha, \beta]$. (If the interval is not bounded, we assume further that $x(t)$ is uniformly continuous). Define $x + y$ and $\alpha x$ by

$$(x + y)(t) = x(t) + y(t)$$
8. Example of a Hilbert space

\((\alpha x)(t) = \alpha x(t)\).

\(C[\alpha, \beta]\) is a Banach space with the norm given by

\[\|x\| = \sup_{t \in [\alpha, \beta]} |x(t)|\]

Converges in this metric is nothing but uniform convergence on the whole space.

2) \(L_p(\alpha, \beta)\). (1 \(\leq p < \infty\)). This is the space of all real or complex valued Lebesgue functions \(f\) on the open interval \((\alpha, \beta)\) for which \(|f(t)|^p\) is Lebesgue summable over \((\alpha, \beta)\); two functions \(f\) and \(g\) which are equal almost everywhere are considered to define the same vector of \(L_p(\alpha, \beta)\). \(L_p(\alpha, \beta)\) is a Banach space with the norm:

\[\|f\| = \left(\int_{\alpha}^{\beta} |f(t)|^p dt\right)^{1/p}\]

The fact that \(\|\|\) thus defined is a norm follows from Minkowski’s inequality; the Riesz-Fischer theorem asserts the completeness of \(L_p\).

3) \(L_\infty(\alpha, \beta)\): This is the space of all measurable (complex valued) functions \(f\) on \((\alpha, \beta)\) which are essentially bounded, i.e., for every \(f \in L_\infty(\alpha, \beta)\) there exists \(a \varphi > 0\) such that \(|f(t)| \leq \varphi\) almost everywhere. Define \(\|f\|\) to be the infimum of such \(\varphi\).

(Here also we identify two functions which are equal almost everywhere).

8 Example of a Hilbert space

\(L_2(\alpha, \beta) : L_2(\alpha, \beta)\) (see example (2) above), is a Hilbert space with the scalar product

\[(f, g) = \int_{\alpha}^{\beta} f(t)\overline{g(t)}dt.\]
9 Completion of a normed linear space

Just as the completeness of the real number field plays a fundamental role in analysis, the completeness of a Banach space will play an essential role in some of our subsequent discussions. If we have an incomplete normed linear space we can always complete it; we can imbed this space in a Banach spaces as an everywhere dense subspace and this Banach spaces is essentially unique. We have, in fact, the

**Theorem.** Let $X_0$ be a normed linear space. Then there exists a complete normed linear space (Banach spaces ) $X$ and a norm preserving isomorphism $T$ of $X_0$ onto a subspace $X'$ of $X$ which is dense in $X$ in the sense of the norm topology. (That $T$ is a norm preserving isomorphism means that $T$ is one-to-one, $T(\alpha x_0 + \beta y_0) = \alpha T(x_0) + \beta T(y_0)$ and $\|x\| = \|T(x)\|$). Such an $X$ is determined uniquely upto a norm preserving isomorphism

**Sketch of the proof:** The proof follows the same idea as that utilized for defining the real numbers from the rational numbers. Let $X$ be the totality of all Cauchy sequences $\{x_n\} \subset X_0$ classified according to the equivalence: $\{x_n\} \sim \{y_n\}$ if and only if $\lim_{n \to \infty} \|x_n - y_n\| = 0$. Denote by $[x_n]$ the class containing $\{x_n\}$.

If $\tilde{x}, \tilde{y} \in X$ and $\tilde{x} = [x_n], \tilde{y} = [y_n]$, define $\tilde{x} + \tilde{y} = [x_n + y_n], \alpha \tilde{x} = [\alpha x_n], \|\tilde{x}\| = \lim_{n \to \infty} \|x_n\|$. These definitions do not depend on the particular representatives for $\tilde{x}, \tilde{y}$ respectively. Finally if $x_0 \in X_0$ defines $T(x_0) = [x_n]$ where each $x_n = x_0$.

10 Additive operators

**Definition.** Let $X$ and $Y$ be linear spaces over $K$. An additive operator from $X$ to $Y$ is a single-valued function $T$ from a subspace $M$ of $X$ into $Y$ such that

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y), \quad x, y \in M, \alpha, \beta \in K.$$  

$M$ is called the domain of $T$ and is denoted by $\mathcal{D}(T)$; the set $\{z | z \in Y$
such that $z = T x$ for some $x \in \mathcal{D}(T)$ is called the range of $T$ and is denoted by $\mathcal{W}(T)$.

If $Y$ is the space of real or complex numbers (according as $X$ is a real or a complex linear space) and $T$ is an additive operator from $X$ to $Y$ we say that $T$ is an additive functional.

Definition. Let $X$ and $Y$ be two normed linear spaces. An additive operator $T$ is said to be continuous at $x_0 \in \mathcal{D}(T)$ if for every sequence $\{x_n\} \subset \mathcal{D}(T)$ with $x_n \to x_0$ we have $T x_n \to T x_0$. An additive operator is said to be continuous (on $\mathcal{D}(T)$) if it is continuous at every point of $\mathcal{D}(T)$. It is easy to see that an additive operator $T$ is continuous on $\mathcal{D}(T)$ if it is continuous at one point $x_0 \in \mathcal{D}(T)$.

Proposition. An additive operator $T : X \to Y$ between two normed linear spaces is continuous if and only if there exists a real number $\varphi > 0$ such that

$$||T x|| \leq \varphi ||x|| \quad \text{for every} \quad x \in \mathcal{D}(T)$$

Proof. The sufficiency of the condition is evident, for if $x_n \to x_0$ then $||T x_n - T x_0|| \leq \varphi ||x_0 - x_n|| \to 0$.

Now assume that $T$ is continuous. If there exists no $\varphi$ as in the proposition, then there exists a sequence $\{x_n\} \subset \mathcal{D}(T)$ such that $||T x_n|| > n ||x_n||$. Since $T(0) = 0$, $x_n \neq 0$. Define $y_n = x_n / \sqrt{n} ||x_n||$. Then $||y_n|| = 1 / \sqrt{n} \to 0$ as $n \to \infty$; as $T$ is continuous $T y_n$ must tend to zero as $n \to \infty$.

But $T y_n = \frac{1}{\sqrt{n} ||x_n||} T x_n$ and $||T y_n|| = \frac{1}{\sqrt{n} ||x_n||} ||T x_n|| > \sqrt{n}$ and so $T y_n$ does not tend to zero. This is a contradiction. $\square$

Let $T$ be an additive operator from a linear space $X$ into a linear space $Y$. $T$ is one-one if and only if $T x = 0$ implies $x = 0$. If $T$ is one-one it has an inverse $T^{-1}$, which is an additive operator from $Y$ into $X$ with domain $\mathcal{W}(T)$, defined by

$$T^{-1} y = x \quad \text{if} \quad y = T x.$$ 

$T^{-1}$ satisfies the relations $T^{-1} T x = x$ for $x \in \mathcal{D}(T)$ and $T T^{-1} y = y$. 
for \( y \in \mathcal{D}(T^{-1}) = \mathcal{W}(T) \). If \( X \) and \( Y \) are normed linear spaces, \( T \) has a continuous inverse if and only if there exists a \( \delta > 0 \) such that \( \|Tx\| \geq \delta \|x\| \) for \( x \in \mathcal{D}(T) \).

The sum of two operators \( T \) and \( S \), with \( \mathcal{D}(T), \mathcal{D}(S) \subset x \) and \( \mathcal{W}(T), \mathcal{W}(S) \subset Y \) is the operator \( (T + S) \), with domain \( \mathcal{D}(T) \cap \mathcal{D}(S) \), defined by:

\[
(T + S)x = Tx + Sx.
\]
Lecture 2

1 Linear operators

**Definition.** An additive operator $T$ from a normed linear space $X$ into a normed linear space $Y$ whose domain $\mathcal{D}(T)$ is the whole space $X$ and which is continuous is called a linear operator from $X$ to $Y$. The norm $\|T\|$ of a linear operator is by definition: $\|T\| = \|T\|_X = \sup_{x \in X, \|x\| \leq 1} \|Tx\|$. If $Y$ is the real or complex numbers (according as $X$ is a real or a complex linear space) the linear operator $T$ is called a linear functional on $X$.

So far we have proved the existence of non-trivial linear functionals. We shall prove the Hahn-Banach extension theorem which will have as a consequence the existence of many linear functionals on a normed linear space.

2 Hahn-Banach lemma

**Definition.** Let $X$ be a linear space (over real or complex numbers). A real valued function $p$ on $X$ will be called a semi-group (or a sub-additive functional) if it satisfies the following conditions:

i) $p(\alpha x) = |\alpha|p(x)$, for each $\alpha \in K$ and $x \in X$.

ii) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$.

Note that these conditions imply that $p(x) \geq 0$ for all $x \in X$. 
3 Lemma (Hahn-Banach)

Let $X$ be a real linear space and $p$ a semi-norm on $X$. Let $M$ be a (real) subspace of $X$ and $f$ a real additive functional on $M$ such that $f(x) \leq p(x)$ for all $x \in M$. Then there exists a real additive functional $F$ on $X$ such that $F$ is an extension of $f$ (i.e., $F(x) = f(x)$ for $x \in M$) and $F(x) \leq p(x)$ for all $x \in X$.

**Proof.** By the application of Zorn’s lemma or transfinite induction, it is enough to prove the lemma when $X$ is spanned by $M$ and an element $x_0 \notin M$, i.e., when $X = \{M, x_0\} = \{x \mid x \in X, x = m + ax_0, m \in M, a \text{ real} , x_0 \notin M\}$. The representation of an element $x \in X$ in the form $x = m + ax_0, (m \in M, a \text{ real})$ is unique. It follows that if, for any real number $c$, we define $F(x) = f(m) + ac$, then $F(x)$ is an additive functional on $X$ which is an extension of $f(x)$. We have now to choose $c$ in such a way that $F(x) \leq p(x), x \in X$, i.e.,

$$f(m) + ac \leq p(m + ax_0).$$

This condition is equivalent to the following two conditions:

$$\begin{cases} f\left(\frac{m}{a}\right) + c \leq p\left(\frac{m}{a} + x_0\right) \quad \text{for } a > 0 \\ f\left(\frac{m}{a}\right) - c \leq p\left(\frac{m}{a} - x_0\right) \quad \text{for } a < 0. \end{cases}$$

To satisfy these conditions, we shall choose $c$ such that

$$f(m') - p(m' - x_0) \leq c \leq p(m'' + x_0) - f(m'')$$

for all $m', m'' \in M$. Such a choice of $c$ is possible since

$$f(m') + f(m'') = f(m' + m'') \leq p(m' + m'')$$

$$= p(m' - x_0 + m'' + x_0)$$
4. Hahn-Banach extension theorem...

\[ \leq p(m' - x_0) + p(m'' + x_0). \]

So \[ f(m') - p(m' - x_0) \leq p(m'' + x_0) - f(m''), m', m'' \in M. \]

So \[ \sup_{m' \in M} \{ f(m') - p(m' - x_0) \} \leq \inf_{m'' \in M} \{ p(m'' + x_0) - f(m'') \} \]

and we can choose for \( c \) any number in between. \( \square \)

4 Hahn-Banach extension theorem for real normed linear spaces

**Theorem.** Let \( X \) be a real normed linear space and \( M \) a real subspace of \( X \). Given a (real) linear functional \( f \) on \( M \), we can extend \( f \) to a (real) linear functional on the whole space \( X \) in such a way that the norm is preserved:

\[ \|F\| = \|F\|_X = \|f\|_M. \]

**Proof.** Take \( p(x) = \|f\|_M \|x\| \) in the Hahn-Banach lemma. We have \( f(x) \leq p(x) \) on \( M \) and \( p(x) \) is subadditive. We then have an additive functional \( F(x) \) on \( X \) which is an extension of \( f \) with \( F(x) \leq \|f\|_M \|x\| \) for all \( x \in X \). Also \( -F(x) = F(-x) \leq \|f\|_M \|x\| = \|f\|_M \|x\|. \) Hence

\[ |F(x)| \leq \|f\|_M \|x\|. \]

This shows that \( F \) is a linear functional on \( X \) and \( \|F\|_X \leq \|f\|_M \). The reverse inequality, \( \|F\|_X \geq \|f\|_M \), is trivial as \( F \) is an extension of \( f \). \( \square \)

5 Hahn-Banach extension theorem for complex normed linear spaces (Bohnenblust-Sobczyk)

**Theorem.** Let \( X \) be a complex normed linear space and \( M \) a (complex) subspace. Given a complex linear functional \( f \) on \( M \) we can extend \( f \) to a complex linear functional \( F \) on \( X \) in such a way that \( \|F\|_X = \|f\|_M \).
Proof. A complex normed linear space becomes a real normed linear space if scalar multiplication is restricted to real numbers and the real and imaginary parts of a complex linear functional are real functionals. If \( f(x) = g(x) + ih(x) \) (\( g(x), h(x) \) real), \( g \) and \( h \) are real linear functionals on \( M \) and \( \|g\|_M \leq \|f\|_M, \|h\|_M \leq \|f\|_M \). Since, for each \( x \in M \),

\[
    g(ix) + ih(ix) = f(ix)
    = if(x)
    = i(g(x) + ih(x))
    = -h(x) + ig(x),
\]

we have \( h(x) = -g(ix) \), for \( x \in M \).

By the Hahn-Banach theorem for real linear spaces \( g \) can be extended to a real linear functional \( G \) on \( X \) with the property \( \|G\|_X = \|g\|_M \). Now define

\[
    F(x) = G(x) - iG(ix).
\]

\( F \) is then a complex linear functional on \( X \). (For complex additivity notice that

\[
    F(ix) = G(ix) - iG(-x) = G(ix) + iG(z) = iF(x).
\]

\( F \) is an extension of \( f \); for, if \( x \in M \),

\[
    F(x) = G(x) - iG(ix) = g(x) - ig(ix) = g(x) + ih(x) = f(x).
\]

We have now only to show that the norm is not changed. For this, writes, for \( x \in X, F(x) = re^{it} \). Then \( e^{-it}F(x) \) is real. So

\[
    |F(x)| = |e^{-it}F(x)| = |F(e^{-it}x)|
    = |G(e^{-it}x)| \quad (= \text{ since } e^{-it}F(x) \text{ is real}).
    \leq \|G\| \|e^{-it}x\|
    = \|g\|_M \|x\|
    \leq \|f\|_M \|x\|.
\]

So \( \|F\|_X \leq \|f\|_M \) and the reverse inequality holds since \( F \) is an extension of \( f \). \qed
6 Existence of non-trivial linear functionals

We consider some consequences of the Hahn-Banach extension theorem; we prove the existence of plenty of linear functionals on a normed linear space.

**Proposition.** Let $X$ be a normed linear (real or complex) and $x_0 \neq 0$ be an elements of $X$. Then there exists a linear functional $f_0$ on $X$ such that $f_0(x_0) = ||x_0||$ and $||f_0|| = 1$.

**Proof.** Let $M$ be the subspace spanned by $x_0$, i.e., $M = \{x | x = \alpha x_0 \text{ for some number } \alpha \}$. Define $f(x) = \alpha ||x_0||$ for $x = \alpha x_0 \in M$. This is a linear functional on $M$ and $||f||_M = 1$. By the Hahn-Banach extension theorem there exists a linear functional $f_0$ on $X$ which extends $f$ in such a way that $||f_0|| = ||f||_M = 1; f_0(x_0) = f(x_0) = ||x_0||$. □

**Remark.** For a pre-Hilbert space the existence of such a linear functional is evident; we may take $f_0(x) = \langle x, \frac{x_0}{||x_0||} \rangle$. The additivity of $f_0$ follows from the homogeneity and distributivity of the scalar product. The continuity of $f_0$ is a consequence of Schwarz’s inequality.

**Proposition.** Let $X$ be a normed linear space. Let $M$ be a subspace and $x_0$ an element $X$ such that $d = \inf_{m \in M} ||x_0 - m|| > 0$. Then there exists a linear functional $f_0$ on $X$ such that $f_0(x) = 0$ for every $x \in M$ and $f_0(x_0) = 1$.

**Proof.** Let $M_0 = \{x | x = m + \alpha x_0, m \in M \}$. Define $f(x) = \alpha$ for $x = m + \alpha x_0 \in M_0(m \in M)$. $f$ is additive on $M_0$, vanishes on $M$ and $f(x_0) = 1$. Also $f$ is continuous on $M_0$: if $\alpha \neq 0$, then $x = m + \alpha x_0 \neq 0(m \in M)$, and

$$|f(x)| = |\alpha| = \alpha||x||/||x|| = ||\alpha||/||m + \alpha x_0|| = \|x\|/||x_0 + (-m/\alpha)|| \leq d^{-1}||x||(-m/\alpha \in M);$$
if $\alpha = 0$, $f(x) = 0$ and the inequality $|f(x)| \leq d^{-1}||x||$ is still valid. If $f_o$ is a linear functional on $X$ which is an extension of $f$, then $f_o$ satisfies the requirements of the proposition.

$\square$

7 Orthogonal projection and the Riesz representation theorem

**Definition.** Let $x$ and $y$ be two elements of a pre-Hilbert space $X$; we say that $x$ is orthogonal to $y$ (written $x \perp y$) if $(x, y) = 0$. If $x \perp y$ then $y \perp x$; if $x \perp x$, then $x = 0$.

Let $M$ be a subset of a pre-Hilbert space; we denote by $M^\perp$ the set of elements $x \in X$ such that $x \perp y$ for every $y \in M$.

**Theorem.** Let $M$ be a closed linear subspace of a Hilbert space $X$. Then any $x_o \in X$ can be decomposed uniquely in the form $x_o = m + n$, $m \in M$, $n \in M^\perp$. ($m$ is called the orthogonal projection of $x_o$ on $M$ and is denoted by $P_M x_o$).

**Proof.** The uniqueness of the decomposition is clear from the fact that an element orthogonal to itself is zero. To prove the existence of the decomposition we may assume $M \neq X$ and $x_o \notin M$ (if $x_o \in M$ we have the trivial decomposition with $n = 0$). Let $d = \inf_{m \in M} ||x_o - m||$; since $M$ is closed and $x_o \notin M, d > 0$. Let $\{m_k\} \subset M$ be a minimizing sequence, i.e., $\lim_{k \to \infty} ||x_o - m_k|| = d$. $\{m_k\}$ is a Cauchy sequence; for

$$||m_k - m_n||^2 = ||(x_o - m_n) - (x_o - m_k)||^2$$

$$= 2(||x_o - m_n||^2 + ||x_o - m_k||^2) - 2||x_o - m_k - m_n||^2$$

$((\text{Euclidean property}))$

$$= 2(||x_o - m_n||^2 + ||x_o - m_k||^2) - 4||x_o - \frac{m_k + m_n}{2}||^2$$

$$\leq 2(||x_o - m_n||^2 + ||x_o - m_k||^2) - 4d^2 (\text{as } \frac{m_k + m_n}{2} \in M)$$

$$\to 2(d^2 + d^2) - 4d^2 = \text{as } m, n \to \infty.$$
7. Orthogonal projection and...

$m \in X$ with $\lim_{k \to \infty} ||m - m_k|| = 0$; in fact $m \in M$, as $M$ is closed. Also $||x_0 - m|| = d$. Write $x_0 = m + (x_0 - m)$. Putting $n = x_0 - m$ we have to show that $n \in M^\perp$. Let $m' \in M$. Since, for any real $\alpha$, $m + \alpha m' \in M$ we have $d^2 \leq ||x_0 - m - \alpha m'||^2 = ||n - \alpha m'||^2 = (n - \alpha m', n - \alpha m')$

$= ||n||^2 - \alpha(n, m') - \alpha(m', n) + \alpha^2||m'||^2$.

Since $||n||^2 = d^2$, this gives, for any real $\alpha$,

$0 \leq -2\alpha \mathcal{R}(n, m') + \alpha^2||m'||^2$.

So $\mathcal{R}(n, m') = 0$ for every $m' \in M$. Replacing $m'$ by $im'$ we have $\mathcal{R} m(n, m') = 0$, for every $m' \in M$. Thus $(n, m') = 0$ for each $m' \in M$.

**Remark**. If $x_0 \notin M$, then $n \neq 0$ and $f_n(x) = (x, \frac{n}{||n||^2})$ satisfies the conditions of the last proposition.

**Theorem Riesz**. Let $X$ be a Hilbert space and $f$ a linear functional on $X$. Then there exists a unique element $y_f$ of $X$ such that

$$f(x) = (x, y_f)$$

for every $x \in X$.

**Proof.** **Uniqueness**: If $(x, y_1) = (x, y_2)$ for every $x$, $(x, y_1 - y_2) = 0$ for every $x$; choosing $x = y_1 - y_2$ we find $y_1 - y_2 = 0$.

**Existence**: Let $M$ be the zero manifold of $f$, i.e., $M = \{x|f(x) = 0\}$. Since $f$ is additive, $M$ is a linear subspace and since $f$ is continuous $M$ is closed. The theorem is evident if $M = X$. i.e., if $f(x) = 0$ on $X$; in this case we need only take $y_f = 0$. So suppose $M \neq X$. Then there exists, by the last theorem, an element $y_0 \neq 0$ such that $y_0$ is orthogonal to every element of $M$. Define

$$y_f = \frac{f(y_0)}{||y_0||^2} y_0.$$
$y_f$ meets the condition of the theorem. First, for $x \in M$, $f(x) = (x, y_f)$ since $f(x) = 0$ for $x \in M$ and $y_f \in M^\perp$. For elements $x$ of the form $x = \alpha y_0$.

$$(x, y_f) = (\alpha y_o, y_f) = \left( \alpha, \frac{f(y_0)}{||y_0||^2} y_0 \right)$$

$$= \alpha f(y_o) = f(\alpha y_o)$$

$$= f(x).$$

Since $f$ is linear and $(x, y_f)$ is linear and $(x, y_f)$ is linear in $x$, to show that $f(x) = (x, y_f)$ for each $x \in X$ it is enough to show that $X$ is spanned by $M$ and $y_o$. If $x \in X$, write, noting that $f(y_f) \neq 0$,

$$x = \frac{f(x)}{f(y_f)} y_f + \left( x - \frac{f(x)}{f(y_f)} y_f \right).$$

$$\frac{f(x)}{f(y_f)} y_f$$ is of the form $\alpha y_o$. The second term is an element of $M$, since

$$f \left( x - \frac{f(x)}{f(y_f)} y_f \right) = f(x) - \frac{f(x)}{f(y_f)} y_f = 0. \quad \square$$

Remark.

$$||f|| = ||y_f||.$$
Lecture 3

1 The Conjugate space (dual) of a normed linear space

Let $X$ be a normed linear space. Let $X^*$ be the totality of all linear functionals on $X$. $X^*$ is a linear space with the operations defined by:

$$(f + g)(x) = f(x) + g(x), \quad f, g \in X^*, \quad x \in X$$

$$(\alpha f)(x) = \alpha f(x), \quad \alpha \in \mathbb{C}, \quad f \in X^*, \quad x \in X$$

$X^*$ is a Banach space with the norm

$$\|f\| = \sup_{\|x\| \leq 1} |f(x)| \quad (f \in X^*, \quad x \in X).$$

We call the Banach space $X^*$ the conjugate space of $X$.

2 The Resonance Theorem

Lemma Gelfand. Let $p(x)$ be a semi-norm on a Banach space $X$. Then there exists a number $\varphi > 0$ such that

$$p(x) \leq \varphi \|x\|$$

for all $x \in X$

if and only if $p(x)$ is lower semi-continuous. (Lower semi-continuity means this): for any $x_o \in$ and any $\mathcal{E} > 0$, there exists a $\delta = \delta(x, \mathcal{E}) > 0$ such that $p(x) \geq p(x_o) - \mathcal{E}$ for $\|x - x_o\| \leq \delta$. 

21
Proof. i) Suppose \( p(x) \leq \varphi \|x\| \) for all \( x \in X \), \( \varphi > 0 \); then

\[
p(x_o) = p(x_o - x + x) \leq p(x_o - x) + p(x) \leq \varphi \|x - x_o\| + p(x) \leq p(x) + \epsilon, \text{ if } \|x - x_0\| \leq E/\varphi = \delta.
\]

ii) Conversely assume that \( p(x) \) is lower semi - continuous.

To prove that there is a \( \varphi > 0 \) such that \( p(x) \leq \varphi \|x\| \) for every \( x \in X \) it is sufficient to show that \( p(x) \) is bounded, say by \( P_1 \), in some closed sphere \( K \) of positive radius \( (K = \{ x \mid \|x - x_o\| \leq \delta \}) \). For if \( x \in X \) with \( \|x\| \leq \delta \), then \( x_o \) and \( x_o + x \) both belong to \( K \) and hence

\[
\begin{align*}
p(x) &= p(-x_o + x_o + x) \\
&= p(x_o) + p(x_o + x) \\
&\leq 2\varphi_1;
\end{align*}
\]

if \( x \) is an arbitrary element of \( X \)

\[
p(x) = p \left( \frac{\|x\|}{\delta} \cdot \frac{\delta}{\|x\|} \right) = \frac{\|x\|}{\delta} p \left( \frac{\delta}{\|x\|} \right) \\
&\leq \frac{2}{\delta} \frac{\varphi}{\|x\|} \left( a \| \frac{\delta}{\|x\|} \right) = \delta \text{ and choose } \varphi = 2\varphi_1/\delta.
\]

Now we assume that \( p(x) \) is unbounded in every closed sphere of positive radius and derive a contradiction. Let

\[
K_o = \{ x \mid \|x - x_o\| \leq \delta, \delta > 0 \};
\]

there exists in interior point \( x_1 \) of \( K_o \) such that \( p(x_1) > 1 \). By the lower semi - continuity of \( p \), there exists a closed sphere \( K_1 = \{ x \mid \|x - x_1\| \leq \delta_1 < 1, \delta_1 > 0 \}, K_1 \subset K_o \) such that \( p(x) > 1 \) for each \( x \in K_1 \). By a repetition of this argument we may choose a sequence of closed spheres \( K_n = \{ x \mid \|x - x_n\| \leq \delta_n < 1/n, \delta_n > 0 \}, n \) running through all positive integers, such that \( K_n \subset K_{n-1} \) and \( p(x) > n \) for each \( x \in K_n \). For \( m, m' > n \), since \( x_m, x_m' \in K_n \), we have

\[
\|x_m - x_m'\| \leq \|x_m - x_n\| + \|x_m' - x_n\| \leq 2n. \]

3. Weak convergence

Let $X$ be a Banach space, $Y_n(n = 1, 2, \ldots)$ a sequence of normed linear spaces, and $T_n$ a sequence of linear operators from $X$ to $Y_n$. Then the boundedness of the sequences $\{ ||T_n x|| \}$ for every $x \in X$ impels the boundedness of the sequence $\{ ||T_n|| \}$.

**Proof.** For each $x \in X$, $\sup_n ||T_n(x)||$ is finite as $\{ ||T_n(x)|| \}$ is bounded. Define $p(x) = \sup_n ||T_n(x)||$; $p(x)$ is a semi-norm on $X$. $p(x)$ is also lower semi-continuous since it is the supremum of the sequence of continuous functions $\{ ||T_n|| \}$. Consequently, by Gelfand’s lemma, $p(x) \leq \varphi ||x||$ (for some $\varphi > 0$) for such $x \in X$; so $||T_n(x)|| \leq \varphi ||x||$ for each $n$ and each $x \in X$. Thus $||T_n|| \leq \varphi$.

**Corollary.** Let $X$ be a Banach space $Y$ a normed linear space, and $\{ T_n \}$ a sequence of linear operators form $X$ to $Y$. Assume that $s - \lim_{n \to \infty} T_n(x) \in Y$ exists for each $x \in X$. If we define $T x = s - \lim_{n \to \infty} T_n(x)$ then $T$ is a linear operator from $X$ to $Y$ and $||T|| \leq \lim_{n \to \infty} ||T_n||$.

$T$ is evidently additive. By the Resonance theorem, $||T_n(x)|| \leq \varphi ||x||$ ($\varphi > 0$); so $||T(x)|| \leq \varphi ||x||$, i.e., $T$ is continuous. Further, $||T_n x|| \leq ||T_n|| ||x||$; so $||T x|| \leq \lim_{n \to \infty} ||T_n|| ||x||$. Hence $||T|| \leq \lim_{n \to \infty} ||T_n||$.

3 Weak convergence

**Definition.** Let $X$ be a normed linear space; we say that a sequence $\{x_n\} \subset X$ converges weakly to $x_\infty \in X$ (and write $w \lim_{n \to \infty} x_n = x_\infty$) if, for every linear functional $f$ on $X$, we have $\lim_{n \to \infty} f(x_n) = f(x_\infty)$. 

2\delta_n < 2/n; so $x_n$ is a Cauchy sequence. Since $X$ is complete there exists an $x_\infty \in X$ such that $s - \lim_{n \to \infty} x_n = x_\infty$. As $||x_m - x_n|| \leq \delta_n$ for $m > n$, we have, passing to the limits, $||x_\infty - x_n|| \leq \delta_n$. So $x_\infty \in \bigcap_{n=1}^\infty K_n$; this would mean that $p(x_\infty)$ (which is a real number) is greater than every positive integer $n$, which is absurd. □
3. Lecture 3

**Proposition.**

i) $\lim_{n \to \infty} x_n$, if it exists, is unique.

ii) $s - \lim x_n = x_\infty$ implies $w - \lim x_n = x_\infty$.

(The converse is not true in general).

iii) if $w - \lim_{n \to \infty} x_n = x_\infty$ then $\lim_{n \to \infty} ||x_n|| \geq x_\infty$.

**Proof.**

(i) Let $w - \lim x_n = x_\infty, w - \lim x_n = x_\infty, \neq x_\infty$. By the Hahn-Banach theorem there exists a linear functional $f$ on $X$ such that $f(x_\infty - x_\infty) \neq 0$ i.e., $f(x_\infty) \neq f(x_\infty')$. But by the condition of weak limit we must have $f(x_\infty) = \lim_{n \to \infty} f(x_n) = f(x_\infty)$.

(ii) This follows form the inequality:

$$|f(x_\infty) - f(x_n)| = f(x_\infty - x_n) \leq ||f|| ||x_\infty - x_n||,$$

for each $f \in X^*$.

(iii) Let $f_\circ \in X^*$ with $||f|| = 1$ and $f_\circ(x_\infty) = ||x_\infty||$.

Then

$$||x_\infty|| = f_\circ(x_\infty) \leq \lim_{n \to \infty} |f_\circ(x_n)|$$

$$\leq \lim_{n \to \infty} ||f_\circ|| ||x_n||$$

$$= \lim_{n \to \infty} ||x_n||.$$
4. A counter-example

Let \( f \in L_2(0, 1) \). By Bessel’s inequality,

\[
\sum_{n=1}^{\infty} \left| \int_{0}^{1} f(t) \sin n\pi t \, dt \right|^2 \leq \int_{0}^{1} |f(t)|^2 \, dt;
\]

so \( \int_{0}^{1} f(t) \sin n\pi t \, dt \to 0 \) as \( n \to \infty \). But \( \{ \sin n\pi t \} \) is not strongly convergent, since

\[
\| \sin n\pi t - \sin m\pi t \|^2 = \int_{0}^{1} |\sin n\pi t - \sin m\pi t|^2 \, dt = 2 \text{ for } n \neq m.
\]
Lecture 4

1 Local weak compactness of a Hilbert space

**Theorem.** Let \( \{x_n\} \) be a bounded sequence of elements of a Hilbert space (i.e., \(||x_n|| \leq C < \infty, n = 1, 2, \ldots \) ); then we can choose a subsequence of \( \{x_n\} \) which converges weakly to an element of X.

**Proof.** Let \( M \) be the closed linear space spanned by \( \{x_n\} \). \( (M \) is the closure in the sense of the norm of the set of all finite linear combinations \( \sum \alpha_i x_i \) of the elements \( \{x_i\} \). \( M \) is separable, there exists a countable set of elements \( \{y_n\} \) which is dense in \( M \). We may take for example, the rational linear combinations of \( \{x_i\} \) if \( X \) is real and if \( X \) is complex, linear combinations of \( \{x_i\} \) with coefficients of the form \( p + iq \), \( p, q \) rational. \( \Box \)

For each \( y_k \) from \( \{y_n\} \) the sequence \( \{(x_n, y_k)\} \) is bounded ; \( |(x_n, y_k)| \leq ||x_n|| ||y_k|| \leq C ||y_k|| \). By the Bolzano - Weierstrass theorem and a diagonal process we can find a subsequence \( \{x'_n\} \) of \( \{x_n\} \) such that \( \{(x'_n, y_k)\} \) converges for every \( k \). Actually \( \{(x'_n, z)\} \) converges for each \( x \in X \). To prove this, let \( z = y + \omega \) where \( y = P_M z, \omega \in M^\perp \). Then \( (x_n, z) = (x_n, y) \) and we have to prove that \( (x_n, y)(y \in M) \) is convergent. We have

\[
|\langle x_n - x_m', y \rangle| = |\langle x_n - x_m', y - y_k + y_k \rangle|
\leq |\langle x_n - x_m', y_k \rangle| + |\langle x_n - x_m', y - y_k \rangle|
\leq |\langle x_n - x_m', y_k \rangle| + ||x_n - x_m', y - y_k||
\leq |\langle x_n - x_m', y_k \rangle| + 2C ||y - y_k||.
\]

Since \( \{(x_n, y_k)\} \) is convergent and \( \{y_k\} \) is dense in \( M \), it follows that
(xₙ, y) is a Cauchy sequence; so \{(xₙ, y)\} is convergent. Define \( g(z) = \lim_{n \to \infty} f(xₙ, z) \) is continuous. By the Riesz theorem there exists and element \( x_∞ \in X \) such that \( f(z) = (z, x_∞) \) for each \( z \in X \). Since \( \lim_{n' \to \infty} (xₙ, z) = (x_∞, z) \) for each \( z \in X \).

We mention without proof that \( L_p(α,β), 1 < p < ∞ \) is locally weakly compact. But \( L(α,β), L_∞(α,β) \) and \( C[α,β] \) are not locally weakly compact.

We next prove a theorem which will be needed in the study of Cauchy’s problem.

2 Lax-Milgram theorem

Let \( B(u, v) \) be a bilinear functional on a real Hilbert space \( X \) such that

(i) there exists a \( ϕ > 0 \) such that \( |B(u, v)| \leq ϕ||u||||v|| \) for all \( u, v \in X \),

(ii) there exists a \( δ > 0 \) such that \( δ||u||^2 \leq B(u, u) \) for each \( u \in X \).

Then there exists a linear operator \( S \) from \( X \) to \( X \) such that

\[
(u, v) = B(u, Sv)
\]

and \( ||S|| \leq δ^{-1} \).

**Proof.** Let \( V \) be the set of elements \( v \) for which there exists an element \( v^* \) such that \( (u, v) = B(u, v^*) \) for all \( u \in X \). (\( V \) is non-empty; \( 0 \in V \).)

\( v^* \) is uniquely determined by \( v \). For, if \( w \in X \) be such that \( B(u, w) = 0 \) for all \( u \), then \( w = 0 \) as \( δ||w||^2 \leq B(w, w) = 0 \) or \( ||w|| = 0 \). \( V \) is a linear subspace. We have an additive operator \( S \) with domain \( V \), defined by \( S v = v^* \). \( S \) is continuous;

\[
δ||S v||^2 \leq B(S v, S v) = (S v, v) \leq ||S v|| ||v||
\]

so that \( ||S v|| \leq δ^{-1}||v|| \) (if \( ||S v|| = 0 \) this is trivially true). Moreover \( V \) is closed subspace of \( X \). For, of \( v_n \in V \) and \( v_n \to v \in X \), then \( S v_n \) is a Cauchy sequence and so has a limit \( v^* \); but \( (u, v_n) \to (u, v) \) and by
(i) $B(u, Sv) \to B(u, v^*)$ so that $(u, v) = B(u, v^*)$ for each $u$; so $v \in V$. The proof will be complete if we show that $V = X$. Suppose $V \neq X$. Then there exists $w \in X$ such that $w \neq 0$ and $(w, v) = 0$ for each $v \in V$. Consider the functional, as

$$|F(z)| = |B(z, w)| \leq \varphi||z|| ||w||.$$

So by Riesz’ theorem, there exists, $w' \in X$ such that $B(z, w) = (z, w')$ for each $z \in X$. So $w' \in V$ and $S w' = w$. So

$$\delta||w||^2 \leq B(w, w) = (w, w')$$

$$= 0,$$

i.e.,

$$w = 0$$

which is a contradiction. □
Part II

Semi-group Theory
Definition. Let \( \{T_t\}_{t \geq 0} \) be a one-parameter family of linear operators on a Banach space \( X \) into itself satisfying the following conditions:

1. \( T_t T_s = T_{t+s}, \ T_0 = I, I \) denoting the identity operator on \( X \) (Semi-group property).
2. \( \lim_{t \to t_0} T_t x = T_{t_0} x \leq 0 \) and each \( x \in X \) (strong continuity).
3. There exists a real number \( \beta \geq 0 \) such that \( \|T_t\| \leq e^{\beta t} \) for \( t \geq 0 \).

We call such a family \( \{T_t\} \) a semi group of linear operators of normal type on the Banach space \( X \), or simply a semi-group.

Remark. The third condition may look a bit curious but it is nothing but a restriction of the order of \( \|T_t\| \) near \( t = 0 \), because we can prove the following.

Proposition. The two conditions (1) and (2) imply the following:

1. \( \lim_{t \to \infty} t^{-1} \log \|T_t\| = \phi < \infty (\phi \ may \ be \ -\infty). \)
2. \( \|T_t\| \) is bounded in any bounded interval \([0, t_0], \ 0 < t_0 < \infty \).

Proof. We first prove (4). Suppose \( \|T_t\| \) is unbounded in some interval \([0, t_0], \ 0 < t_0 < \infty \). Then there would exist a sequence \( \{t_n\} \) \( (n = 11, 2, \ldots) \) such that \( \|T_{t_n}\| \geq n \) and \( 0 < \lim_{n \to \infty} t_n = t_\infty < t_0 < \infty \). Since \( \|T_{t_\infty}\| \) is unbounded, by the resonance theorem, \( \|T_{t_\infty} x\| \) is unbounded at least for one \( x \in X \); but by strong continuity, \( \lim_{n \to \infty} T_{t_n} x = T_{t_\infty} x \) for each \( x \in X \). This is a contradiction.

To prove (3'), let \( p(t) = \log \|T_t\|, \ p(t) < \infty \) (may be \( -\infty \)). Since \( \|T_{t+s}\| = \|T_t T_s\| \leq \|T_t\| \|T_s\| \), we have \( p(t+s) \leq p(t) + p(s) \). Let \( \phi \inf_{t>0} r^{-1} p(t) \) is either finite or \( -\infty \). We shall show that \( \lim_{t \to \infty} t^{-1} p(t) \) exists and is equal to \( \phi \). Assume, first, \( \phi \) is finite. Choose for any \( \epsilon > 0 \), a number \( a > 0 \) in such a way that \( p(a) \leq (\phi + \epsilon)a \). Let \( n \) be an integer such that \( na \leq t < (n+1)a \).

Then

\[
\phi \leq \frac{p(t)}{t} \leq \frac{p(na)}{t} + \frac{p(t-na)}{t}
\]
\[
\begin{align*}
&\leq \frac{na}{t} p(a) + \frac{p(t - na)}{t} \\
&\leq \frac{na}{t} (\varphi + E) + \frac{p(t - na)}{t}.
\end{align*}
\]

Letting \( t \to \infty \), \( \frac{p(t - na)}{t} \) tends to zero since \( p(t - na) \) is bounded from above (since, as we have proved above, \( \|T_s\| \) is bounded in any finite interval of \( s \)). Thus \( \lim_{t \to \infty} t^{-1} p(t) = \varphi \). The case \( \varphi = -\infty \) can be treated similarly. \( \square \)
Lecture 5

1 Some examples of semi-groups

I In $C[0, \infty]$ the space of bounded uniformly continuous functions on the closed interval $[0, \infty]$ define $\{T_t\}_{t \geq 0}$ by

$$(T_t x)(s) = x(t + s) \ (x \in C).$$

$\{T_t\}$ is a semi-group. Condition (1) is trivially verified. (2) follows from the uniform continuity of $x$, as

$$\|T_t x - T_{t_0} x\| = \sup_{s \geq 0} |x(t + s) - x(t_0 + s)|.$$

Finally $\|T_t\| = 1$ and so (3) is satisfied with $\beta = 0$.

In this example, we could replace $C[0, \infty]$ by $C[-\infty, \infty]$.

II On the space $C[0, \infty]$ (or $C[-\infty, \infty]$), define $\{T_t\}_{t \geq 0}$

$$(T_t x)(s) = e^{\beta t} x(s)$$

where $\beta$ is a fixed non-negative number. Again (1) is trivial; for (2) we have $\|T_t x - T_{t_0} x\| = |e^{\beta t} - e^{\beta t_0}| \sup_s |x(s)|$. Trivially $\|T_t\| = e^{\beta t}$.

III Consider the space $C[-\infty, \infty]$. Let

$$N_t(u) = \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t}, \quad -\infty < u < \infty, \ t > 0,$$
(the normal probability density). Define \( \{T_t\}_{t \geq 0} \) on \( C[-\infty, \infty] \) by:

\[
(T_t x)(s) = \begin{cases} 
\int_{-\infty}^{\infty} N_t(s-u)x(u)du, & \text{for } t > 0 \\
x(s) & \text{for } t = 0
\end{cases}
\]

Each \( T_t \) is continuous:

\[
\|T_t x\| \leq \|x\| \int_{-\infty}^{\infty} N_t(s-u)du = \|x\|, \quad \text{as } \int_{-\infty}^{\infty} N_t(s-u)du = 1.
\]

Moreover it follows from this that condition (3) is valid with \( \beta = 0 \). By definition \( T_{t_0} = I \) and the semi-group property \( T_t T_s = T_{t+s} \) is a consequence of the well-known formula concerning the Gaussian distribution.

\[
\frac{1}{\sqrt{2\pi(t+t')}} e^{-u^2/2(t+t')} = \frac{1}{\sqrt{2\pi t'}} \int_{-\infty}^{\infty} e^{-z^2/2t} \frac{e^{-z^2/2t'}}{\sqrt{2\pi t'}} dv.
\]

(Apply Fubini’s theorem). To prove the strong continuity, consider \( t, t_0 > 0 \) with \( t \neq t_0 \). (The case \( t_0 = 0 \) is treated in a similar fashion). By definition

\[
(T_t x)(s) - (T_{t_0} x)(s) = \int_{-\infty}^{\infty} \left\{ N_t(s-u)x(u) - N_{t_0}(s-u)x(u) \right\} du.
\]

The integral \( \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-(s-u)^2/2t} x(u)du \) becomes, by the change of variable \( \frac{s-u}{\sqrt{t}} = z \),

\[
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2t} x(s - \sqrt{t}z)dz. \quad \text{Hence}
\]

\[
(T_t x)(s) - (T_{t_0} x)(s) = \int_{-\infty}^{\infty} N_1(z) \left\{ x(-s\sqrt{t}) - x(s - \sqrt{t_0}z) \right\} dz \cdot x(s)
\]
being uniformly continuous on \(-\infty, \infty\), for any \(\varepsilon > 0\) there exists a number \(\delta = \delta(\varepsilon) > 0\) such that \(|x(s_1) - x(s_2)| \leq \varepsilon\) whenever \(|s_1 - s_2| \leq \delta\). Now, splitting the last integral

\[
|(T_t x)(s) - (T_{t_0} x)(s)| \\
\leq \int_{|\sqrt{t} - \sqrt{t_0}| \leq \delta} N_1(z) |x(s - \sqrt{t} z) - x(s - \sqrt{t_0} z)| dz \\
+ \int_{|\sqrt{t} - \sqrt{t_0}| > \delta} N_1(z)(\ldots) dz \\
\leq \mathcal{E} \int N_1(z) dz + 2||x|| \int_{|\sqrt{t} - \sqrt{t_0}| > \delta} N_1(z) dz \\
= \mathcal{E} + 2||x|| \int_{|z| > |\sqrt{t} - \sqrt{t_0}|} N_1(z) dz
\]

The second term on the right tends to 0 as \(|t - t_0| \to 0\), because the integral \(\int_{-\infty}^{\infty} N_1(z) dz\) converges. Thus

\[
\lim_{t \to t_0, -\infty < s < \infty} \sup_{-\infty < t < \infty} |(T_t x)(s) - (T_{t_0} x)(s)| \leq \mathcal{E}.
\]

Since \(\mathcal{E} > 0\) was arbitrary, we have proved the strong continuity at \(t = t_0\) of \(T_t\).

In this example we can also replace \(C[0, \infty]\) by \(L_p[0, \infty]\), \(1 \leq p < \infty\). Consider, for example \(L_1[0, \infty]\). In this case, \(||T_t x|| \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} N_t(s - u)|x(u)| ds du \leq ||x||\), applying Fubini’s theorem.

As for the strong continuity, we have

\[
(T_t x)(s) - (T_{t_0} x)(s)
\]
\begin{align*}
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} N_1(z) \left| x(s - \sqrt{tz}) - x(s - \sqrt{t_0z}) \right| dz ds \\
&\leq \int_{-\infty}^{\infty} N_1(z) \left[ \int_{-\infty}^{\infty} |x(s - \sqrt{tz}) - x(s - \sqrt{t_0z})| ds \right] dz
\end{align*}

Since \( N_1(z) \int_{-\infty}^{\infty} |x(s - \sqrt{tz}) - x(s - \sqrt{t_0z})| ds \leq 2\|x\| N_1(z) \), we may apply Lebesgue’s dominated convergence theorem. We then have

\[ \lim_{t \to t_0} \left( \int_{-\infty}^{\infty} N_1(z) \left\{ \int_{-\infty}^{\infty} |x(s - \sqrt{tz}) - x(s - \sqrt{t_0z})| ds \right\} dz \right) = 0, \]

by the continuity in mean of the Lebesgue integral.

IV Consider \( C[-\infty, \infty] \). Let \( \lambda > 0, \mu > 0 \). Define \( \{ T_t \}_{t \geq 0} \)

\[ (T_t x)(s) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} x(s - k\mu). \]

\( \{ T_t \} \) is a semi-group. Strong continuity follows from:

\[ |(T_t x)(s) - (T_{t_0} x)(s)| \leq \|x\| \left| e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} - e^{-\lambda t_0} \sum_{k=0}^{\infty} \frac{(\lambda t_0)^k}{k!} \right| = 0. \]

(3) is satisfied with \( \beta = 0 \). To verify (1)

\[ (T_w(t) x)(s) = e^{\lambda w} \sum_{l=0}^{\infty} \frac{(\lambda w)^l}{l!} \left[ e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} f(s - k\mu - 1\mu) \right] \]

\[ = e^{-\lambda(w+t)} \sum_{p=0}^{\infty} \frac{1}{p!} \left[ \sum_{l=0}^{p} \frac{(\lambda w)^l (\lambda t)^{p-l}}{l! p^{p-l}} f(s - p\mu) \right] \]

\[ = e^{-\lambda(w+t)} \sum_{p=0}^{\infty} \frac{1}{p!} (\lambda w + \lambda t)^p f(\lambda + \lambda t)^{p} f(s - p\mu) \]
2. The infinitesimal generator of a semi-group

\[ = (T_{w+r}x)(s). \]

2 The infinitesimal generator of a semi-group

Definition. The infinitesimal generator \( A \) of a semi-group \( T_t \) is defined by:

\[ Ax = s - \lim_{h \to 0} h^{-1}(T_h - I)x, \]

i.e., as the additive operator \( A \) whose domain is the set

\[ D(A) = \{ x \mid s - \lim_{h \to 0} h^{-1}(T_h - I)x \text{ exists} \} \]

and for \( x \in D(A) \),

\[ Ax = s - \lim_{h \to 0} h^{-1}(T_h - I)x. \]

\( D(A) \) is evidently non-empty; it contains at least zero. Actually \( D(A) \) is larger. We prove the

Proposition. \( D(A) \) is dense in \( X \) (in the norm topology).  

Proof. Let \( \varphi_n(s) = ne^{-ns} \). Introduce the linear operator \( C_{\varphi_n} \) defined by

\[ C_{\varphi_n}x = \int_0^\infty \varphi_n(s)T_sxd\sigma \text{ for } x \in X \text{ and } n > \beta, \]

the integral being taken in the sense of Riemann. (The ordinary procedure of defining the Riemann integral of a real or complex valued functions can be extended to a function with values in a Banach space, using the norm instead of absolute value). The convergence of the integral is a consequence of the strong continuity of \( T_s \) in \( s \) and the inequality,

\[ \| \varphi_n(s)T_sx \| \leq ne^{(-n+\beta)s} \| x \|. \]

\[ \square \]
The operator $C_{\varphi_n}$ is a linear operator whose norm satisfies the inequality
\[ \| \varphi_n \| \leq n \int_0^\infty e^{(-n+\beta)s} ds = 1/1 - \beta/n. \]

We shall now show that $W(C_{\varphi_n}) \subseteq \mathcal{D}(A)$ (denotes the range of $C_{\varphi_n}$) for each $n > \beta$ and that for each $x \in X$, $s \lim_{n \to \infty} C_{\varphi_n} x = x$; then $\bigcup_{n > \beta} W(C_{\varphi_n})$ will be dense in $X$ and a-portion $\mathcal{D}(A)$ will be dense in $X$. We have
\[
h^{-1}(T_h - I)C_{\varphi_n} x = h^{-1} \int_0^\infty \varphi_n(s)T_hT_s x ds - h^{-1} \int_0^\infty \varphi_n(s)T_s x ds
\]
(The change of the order $T_h \int \cdot = \int T_h \cdot \cdot$ is justified, using the additivity and the continuity of $T_h$, by approximating the integral by Riemann sums). Then
\[
h^{-1}(T_h - I)C_{\varphi_n} x = h^{-1} \int_0^\infty \varphi_n(s)T_{h+s} x ds - h^{-1} \int_0^\infty \varphi_n(s)T_s x ds
\]
\[
= h^{-1} \int_0^\infty \varphi_n(s-h)T_s x ds - h^{-1} \int_0^\infty \varphi_n(s)T_s x ds
\]
( by a change of variable in the first integral ).
\[
= h^{-1} \int_h^\infty (\varphi_n(s-h) - \varphi_n(s))T_s x ds
\]
\[
= h^{-1} \int_0^h \varphi_n(s)T_s x ds.
\]
By the strong continuity of $\varphi_n(s)T_s x$ in $s$, the second term on the
right converges strongly to $-\varphi_n(0)T_0x = -nx$, as $h \downarrow 0$.

$$h^{-1} \int_0^\infty \{\varphi_n(s-h) - \varphi_n(s)\} T_s x ds$$

$$= \int_0^\infty -\varphi'_n(s-\Theta h) T_s x ds \quad (0 < \Theta < 1) \quad \text{(by the mean value theorem)}$$

$$= \int_0^\infty -\varphi'_n(s) T_s x ds + \int_0^h \varphi'_n(s) T_s x ds + \int_h^\infty \varphi'_n(s) - \varphi'_n(s-\theta h) T_s x ds.$$

But, $\int_0^h \varphi'_n(s) T_s x ds \rightarrow 0$ as $h \downarrow 0$ and

$$\left\| \int_0^\infty \{\varphi'_n(s) - \varphi'_n(s-\Theta h)\} T_s x ds \right\|$$

$$\leq n^2 \int_0^\infty \left| e^{-n(s-\theta h)} - e^{-ns} \right| e^{\beta s} \left\| x \right\| ds$$

$$\leq n^2 (e^{\theta h} - 1) \int_0^\infty e^{(\beta - n)s} \left\| x \right\| ds \rightarrow 0 \quad \text{as} \quad h \downarrow 0. (\beta < n).$$

Thus we have proved that $\mathfrak{B}(C_{\varphi_n}) \subseteq \mathfrak{D}(A)$ and

$$AC_{-\varphi_n}x = n(C_{-\varphi_n} - I)x$$

as $\varphi'_n = -n\varphi_n$. Next, we show that $s - \lim_{n \to \infty} C_{\varphi_n}(x) = x$ for each $x \in X$.

We observe that

$$C_{\varphi_n} x - x = \int_0^\infty ne^{-ns} T_s x ds - \int_0^\infty ne^{-ns} x ds, \quad \text{(as} \quad \int_0^\infty ne^{-ns} ds = 1)$$
\[ = n \int_{0}^{\infty} e^{-ns} [T_s x - x] ds. \]

Approximating the integral by Riemann sums and using the triangle inequality we have

\[
\| C_{\varphi_n} x - x \| \leq n \int_{0}^{\infty} e^{-ns} \| T_s x - x \| ds
\]

\[
= n \int_{0}^{\delta} \cdots + n \int_{\delta}^{\delta} \cdots, \delta > 0
\]

\[
= I_1 + I_2, \text{ say.}
\]

Given \( \mathcal{E} > 0 \), by strong continuity, we can choose a \( \delta > 0 \) such that

\[ \| T_s x - x \| < \mathcal{E} \text{ for } 0 \leq s \leq \delta; \text{ then} \]

\[ I_1 \leq \mathcal{E} n \int_{0}^{\delta} e^{-ns} ds \leq \mathcal{E} n \int_{0}^{\infty} e^{-ns} ds = \mathcal{E}. \]

For a fixed \( \delta > 0 \), using the majorization condition in the definition of a semi-group,

\[ I_2 \leq n \int_{\delta}^{\infty} e^{-ns} (e^{\beta s} + 1) \| x \| ds = \| x \| \left[ n \frac{e^{(n+\beta)s\delta}}{-n} \right]_{\delta}^{\infty} - \| x \| \left[ n \frac{e^{-ns\delta}}{-n} \right]_{\delta}^{\infty}. \]

Each of the terms on the right tends to zero as \( n \to \infty \). So \( I_2 \leq \mathcal{E} \), for \( n > n_0 \). Thus \( C_{\varphi_n} x \to x \) as \( n \to \infty \).

**Remark.** That \( \mathcal{D}(A) \) is dense in \( X \) can be proved more easily. But we need the considerations given in the above proof for later purpose.

**Definition.** For \( x \in X \) define \( D_1 T_s x \) by

\[ D_1 T_s x = s - \lim_{h \to 0} h^{-1} (T_{s+h} - T_s)x \]

if the limit exists.
2. The infinitesimal generator of a semi-group

Proposition. If \( x \in \mathcal{D}(A) \) then \( x \in \mathcal{D}(D_t) \) and \( D_t T_t x = AT_t x = T_t Ax \).

Proof. If \( x \in \mathcal{D}(A) \), we have, since \( T_t \) is a linear operator,

\[
T_t Ax = T_t \left( s - \lim_{h \downarrow 0} h^{-1}(T_h - I)x \right)
= s - \lim_{h \downarrow 0} h^{-1}(T_h T_t - T_t)x
= s - \lim_{h \downarrow 0} h^{-1}(T_{t+h} - T_t)x
= s - \lim_{h \downarrow 0} h^{-1}(T_h - I)T_t x = AT_t x.
\]

Thus, if \( x \in \mathcal{D}(A) \), then \( T_t x \in \mathcal{D}(A) \), and

\[
T_t Ax = AT_t x = s - \lim_{h \downarrow 0} h^{-1}(T_{t+h} - T_t)x.
\]

We have now proved that the strong right derivative of \( T_t x \) exists for each \( x \in \mathcal{D}(A) \). We shall now show that the strong left derivative exists and is equal to the right derivative. For this, take any \( f \in X^* \). For fixed \( x \), \( f(T_t x) \) is a continuous numerical function (real or complex-valued) on \( t \geq 0 \). By the above, \( f(T_t x) \) has right derivative \( \frac{d^+ f(T_t x)}{dt} \) and

\[
\frac{d^+ f(T_t x)}{dt} = f(AT_t x) = f(T_t A x).
\]

But \( f(T_t A x) \) is a continuous function. It is well-known that if one of the Dini-derivatives of a numerical function is (finite and) continuous, then the function is differentiable (and the derivative, of course, is continuous). So \( f(T_t x) \) is differentiable in \( t \) and

\[
f(T_t x - x) = f(T_t x) - f(T_0 x)
= \int_0^t \frac{d^+ f(T_s x)}{ds} ds
= \int_0^t f(T_s Ax) ds
= f \left( \int_0^t A x ds \right).
\]
However, if every linear functional vanishes on an element \( x \in X \), then \( x = 0 \) (by Hahn-Banach theorem). Consequently,

\[
T_t x - x = \int_0^t T_s A x \, ds.
\]

for each \( x \in \mathcal{D}(A) \). Since \( T_s \) is strongly continuous in \( s \), it follows from this, that \( T_t \) is strongly derivable:

\[
D_t T_t x = s \lim_{h \to 0} h^{-1} (T_{t+h} - T_t)x
= s \lim_{h \to 0} h^{-1} \int_t^{t+h} T_s A x \, ds
= T_t A x.
\]
Lecture 6

Theorem. For \( n > \beta \), the operator \((I - n^{-1}A)^{-1}\) admits of an inverse \( J_n = (I - n^{-1}A)^{-1} \) which is linear and satisfies the relation

\[
J_n x = n \int_0^\infty e^{-ns} T_s x ds, \quad \text{for } x \in X \quad \text{(i.e., } J_n = C\varphi \text{). Also } \| J_n \| \leq (1 - n^{-1} \beta)^{-1}.
\]

Proof. We first show that \((I - n^{-1}A)^{-1}\) exists [i.e., \((I - n^{-1}A)\) is one-one]. If \((I - n^{-1}A)\) is not one-one, there will exist \( x_0 \in \mathcal{D}(A) \) such that \( \| x_0 \| = 1 \) and \((I - n^{-1}A)x_0 = 0\), i.e., \( Ax_0 = nx_0 \). Let \( f_0 \) be a linear functional on \( X \) such that \( \| f_0 \| = 1 \) and \( f_0(x_0) = 1 \). Define \( \varphi(t) = f_0(T_tx_0) = 1 \). Define \( \varphi(t) = f_0(T_tx_0) \). Since \( x_0 \in \mathcal{D}(A) \), \( \varphi(t) \) is differentiable and

\[
\frac{d\varphi(t)}{dt} = f_0(D_t T_tx_0) = f_0(T_t Ax_0) = f_0(T_t nx_0) = n f_0(T_t x_0) = n \varphi(t).
\]

Solving this differential equation with the initial condition \( \varphi(0) = 1 \) we get \( \varphi(t) = e^{nt} \). On the other hand we have

\[
|\varphi(t)| = |f_0(T_t x_0)| \leq \| f_0 \| \| T_t \| \| x_0 \| \leq e^{\beta t};
\]

since \( \varphi(t) = e^{nt} \) and \( n > \beta \) this is impossible. So \((I - n^{-1}A)^{-1}\) exists.
Since $A C_{\psi_n} x = n(C_{\psi_n} - I)x$, we have $(I_n - n^{-1}A)C_{\psi_n} x = x$ for all $x \in X$. So $(I - n^{-1}A)$ maps $\mathcal{D}(C_n) \subseteq \mathcal{D}(A)$ on to $X$; thus $(I - n^{-1}A)$ maps $\mathcal{D}(A)$ in a one-one manner onto $X$. It follows that $M(C_{\psi_n}) \mathcal{D}(A)$ and $(I - n^{-1}A)^{-1} = C_{\psi_n}$. But $C_{\psi_n}$ is a linear operator and we have already proved that $\|C_{\psi_n}\| \leq (1 - n^{-1}\beta)^{-1}$.

Corollary.

$\mathcal{R}(C_{\psi_n}) = \mathcal{D}(A)$

$A J_n x = n(J_n - I)x, \ x \in X.$

$A J_n x = J_n Ax = n(J_n - I)x, \ x \in \mathcal{D}(A)$

$s - \lim_{n \to \infty} J_n x = x, \ x \in X,$

$D_t T_t x = s - \lim_{h \to 0} h^{-1}(T_{t+h} - T_t)x = AT_t x = T_t Ax, \ x \in \mathcal{D}(A).$

1 The resolvent set and the spectrum of an additive operator on a Banach space

We may state our theorem in the terminology of spectral theory.

Let $A$ be an additive operator (with domain $\mathcal{D}(A)$) from a Banach space $X$ into $X$. Let $\lambda$ be a complex number (it is assumed to be real if $X$ is a real space). Regarding the inverse of the additive operator $(\lambda I - A)$ there are various possibilities.

(1) $(\lambda I - A)$ does not admit of an inverse, i.e., there exists an $x \neq 0$ such that $Ax = \lambda x$. We then call $\lambda$ an eigenvalue of $A$ and $x$ an eigenvector belonging to the eigenvalue $\lambda$. In this case we also say that $\lambda$ is in the point-spectrum of $A$.

(2) When $(\lambda I - A)^{-1}$ exists there are three possibilities:

(i) $\mathcal{D}((\lambda I - A)^{-1})$ is not dense in $X$. Then $\lambda$ is said to be in the residual spectrum of $A$.

(ii) $\mathcal{D}((\lambda I - A)^{-1})$ is dense in $X$ but $(\lambda I - A)^{-1}$ is not continuous.

39 In this case $\lambda$ is said to be in the continuous spectrum.
2. Examples

(iii) $\mathcal{D}((\lambda I - A)^{-1})$ is dense in $X$ and $(\lambda I - A)^{-1}$ is continuous in $\mathcal{D}((\lambda I - A)^{-1})$. Then $(\lambda I - A)^{-1}$ can be extended uniquely to a linear operator on the whole space $X$. In this case $\lambda$ is said to be in the \textit{resolvent set}; the inverse $(\lambda I - A)^{-1}$ is called the \textit{resolvent}.

The complement of the resolvent set in the complex plane (or in the real line if $X$ is real) is called the spectrum of $A$.

The first part of the theorem proved above says that if $\{T_t\}$ is a semi-group of normal type ($\|T_t\| \leq e^{\beta t}$) any number $\lambda > \beta$ is in the resolvent set of the infinitesimal generator $A$.

2 Examples

Using these results we now determine the infinitesimal generators of the semi-groups we considered earlier.

$I : C[0, \infty] : (T_t x)(s) = x(t + s)$

Writing $y_n(s) = (J_n x)(s)$ we have

$$y_n(s) = \int_0^\infty e^{-nt} x(t + s) dt$$

$$= \int_s^\infty e^{-n(t-s)} x(t) dt :$$

$$y'_n(s) = -ne^{-n(s-s)} x(s) + n^2 \int_s^\infty e^{-n(t-s)} x(t) dt$$

$$= nx(s) + ny_n(s)$$

Comparing this with the general formula

$$(AJ_n x)(s) = n((J_n - I)x)(s)$$

or

$$Ay_n(s) = ny_n(s) - nx(s)$$
we have \( Ay_n(s) = y'_n(s) \).

For \( n > \beta, \mathfrak{B}(J_n) = \mathcal{D}(A) \). So if \( y \in \mathcal{D}(A) \), \( y'(s) \) exists and belongs to \( C[0, \infty] \) and

\[
(Ay)(s) = y'(s).
\]

Conversely let \( y(s) \) and \( y'(s) \) both belong to \( C[0, \infty] \); we shall show that \( y \in \mathcal{D}(A) \) and \((Ay)(s) = y'(s)\). For define \( x(s) \) by

\[
y'(s) - ny(s) = -nx(s).
\]

Putting \((J_n x)(s) = y_n(s)\), we have, as shown above,

\[
y'_n(s) - ny_n(s) = -nx(s).
\]

Writing \( \omega(s) = y(s) - y_n(s) \), we obtain

\[
\omega'(s) - n\omega(s) = 0
\]

or \( \omega(s) = Ce^{nt} \). But \( \omega(s) \in C[0, \infty] \) and this is possible only if \( C = 0 \). Hence \( y(s) = y_n(s) \in \mathcal{D}(A) \) and so \( (Ay)(s) = y'(s) \). Thus the domain of the infinitesimal generator \( A \) is precisely the set of functions \( y \in C[0, \infty] \) and for such a function \( Ay = y' \). We have thus characterized the differential operator \( \frac{d}{dt} \) as the infinitesimal generator of the semigroup associated with the translation by \( t \).

II. In this we give the characterization of the second derivation as the infinitesimal generator of the semi-group associated with the Gaussian distribution. The space is \( C[-\infty, \infty] \) and

\[
(T_t x)(s) = \begin{cases} 
\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(s-v)^2/2t} x(v) \, dv & \text{if } t > 0 \\
x(s) & \text{if } t = 0.
\end{cases}
\]

We have

\[
y_n(s) = (J_n x)_s = \int_{-\infty}^{\infty} x(v) \left\{ \int_{0}^{\infty} \frac{n}{\sqrt{2\pi t}} e^{-nt-(s-v)^2/2t} \, dt \right\} \, dv
\]
2. Examples

\[ = \int_{-\infty}^{\infty} x(v) \left\{ \int_{0}^{\infty} \frac{2\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{(s-v)^2}{2n\sigma^2}} d\sigma \right\} dv \]

Assuming for moment the formula

\[ \int_{0}^{\infty} e^{-(\sigma^2+c/\sigma^2)} d\sigma = \frac{\sqrt{\pi}}{2} e^{-2c}, \quad c > 0, \quad \text{with} \quad c = \sqrt{n} \frac{|s-v|}{\sqrt{2}}. \]

we get

\[ y_n(s) = \int_{-\infty}^{\infty} x(v) \left( \frac{\sqrt{n}}{2} e^{-\sqrt{2n}|s-v|} \right) dv \]

\[ = \frac{\sqrt{n}}{2} \int_{-\infty}^{\infty} x(v) e^{-\sqrt{2n}|s-v|} = \frac{\sqrt{2}}{2} \left( \int_{-\infty}^{s} + \int_{s}^{\infty} \right) \]

\[ \]

\[ x(v) \]

being continuous we can differentiate twice and we then obtain

\[ y'_n(s) = n \left\{ \int_{s}^{\infty} x(v) e^{-\sqrt{2n}(v-s)} dv - \int_{-\infty}^{s} x(v) e^{-\sqrt{2n}(s-v)} dv \right\} \]

\[ y''_n(s) = n \left\{ -x(s) - x(s) + 2 \sqrt{n} \int_{-\infty}^{s} x(v) e^{-\sqrt{2n}(v-s)} dv \right. \]

\[ + \sqrt{2n} \int_{-\infty}^{s} x(v) e^{-\sqrt{2n}(s-v)} dv \right\} \]

\[ = -2nx(s) + 2ny_n(s). \]

Comparing this with the general formula

\[ (Ay_n)(s) = (AJ_n x)(s) = n \left\{ (J_n - 1)x \right\}(s) \]

\[ = n(y_n(s) - x(s)) \]
we find that $A y_n(s) = \frac{1}{2} y''(s)$. For $n > \beta$, $\mathfrak{M}(J_n) = \mathfrak{D}(A)$. Thus if $y \in \mathfrak{D}(A)$, $y''(s)$ exists and belongs to $C[-\infty, \infty]$ and further $(Ay)(s) = \frac{1}{2} y''(s)$. Conversely, let $y(s)$ and $y''(s)$ both belong to $C[-\infty, \infty]$. Define $x(s)$ by

$$y''(s) - 2ny(s) = -2nx(s).$$

Putting $y_n(s) = (J_n x)(s)$, we have, as shown above,

$$y''_n(s) - 2ny_n(s) = -2nx(s).$$

So, if $\omega(s) = y_n(s) - y(s)$,

$$\omega''(s) - 2n\omega(s) = 0.$$ 

This $\omega(s) = C_1 e^{\sqrt{2}ns} + C_2 e^{-\sqrt{2}ns}$. 

This function cannot be bounded unless both $C_1$ and $C_2$ are zero. Hence $y(s) = y_n(s)$. So $y(s) \in \mathfrak{D}(A)$ and $(Ay)(s) = \frac{1}{2} y''(s)$.

Thus the differential operator $\frac{1}{2} \frac{d^2}{dt^2}$ is the infinitesimal generator of the semi-group associated with the Gaussian process.

We now prove the formula

$$\int_0^\infty e^{-(\sigma^2 + c^2)/2}d\sigma = \sqrt{\pi}/2e^{-2c}, \quad c > 0.$$ 

We start with the formula

$$\int_0^\infty e^{-x^2}dx = \sqrt{\pi}/2.$$ 

Putting $x = \sigma - c/\sigma$, we have

$$\frac{\sqrt{\pi}}{2} = \int_{\sqrt{c}}^\infty e^{-(\sigma - c/\sigma)^2}(1 + c/\sigma^2)d\sigma.$$
2. Examples

\[ e^{2c} \int_{\sqrt{c}}^{\infty} e^{-\left(\sigma^2 + c^2/\sigma^2\right)/(1 + c/\sigma^2)} d\sigma \]

\[ = e^{2c} \left\{ \int_{\sqrt{c}}^{\infty} e^{-\left(\sigma^2 + c^2/\sigma^2\right)} d\sigma + \int_{\sqrt{c}}^{\infty} e^{-\left(\sigma^2 + c^2/\sigma^2\right)c/\sigma^2} d\sigma \right\} \]

Setting \( \sigma = c/t \) in the last integral

\[ \frac{\sqrt{\pi}}{2} = e^{2c} \left( \int_{\sqrt{c}}^{\infty} e^{-\left(\sigma^2 + c^2\sigma^2\right)} d\sigma \right) \]

\[ = e^{2c} \int_{0}^{\infty} e^{\left(\sigma^2 + c^2/\sigma^2\right)} d\sigma. \]
Lecture 7

1 The exponential of a linear operator

Example III. In $C[-\infty, \infty]$ consider the semi-group associated with a Poisson process, viz.,

$$ (T_t x)(s) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} x(s - k\mu) \lambda, \mu > 0 $$

Since $e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = 1$, we have

$$ \frac{(T_t x)(s) - x(s)}{t} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} (x(s - k\mu) - x(s)) $$

$$ = e^{-\lambda t} \frac{x(s - k\mu) - x(s)}{t} + e^{-\lambda t} \sum_{k=2}^{\infty} \frac{(\lambda t)^k}{k!} (x(s - k\mu) - x(s)). $$

As $t \downarrow 0$ the first term on the right tends uniformly with respect to $s$ to $\lambda(x(s - \mu) - x(s))$; the absolute value of the second term is majorized by $2 \| x \| \frac{e^{-\lambda t}}{t} \sum_{k=2}^{\infty} \frac{(\lambda t)^k}{k!}$ which tends to zero as $t \downarrow 0$. Thus for any $x \in C[-\infty, \infty]$, we have $Ax = \lambda(x(s - \mu) - x(s))$. So in this case the infinitesimal generator is the linear operator defined by:

$$ (Ax)(s) = \lambda[x(s - \mu) - x(s)], $$
for $x \in C[-\infty, \infty]$.

This is the difference generator.

We now intend to represent the original semi-group $\{T_t\}$ by its infinitesimal generator. We expect, by analogy with the case of the ordinary exponential function, the result to be given by

$$T_t x = \exp(tA)x.$$ 

But in general $A$ is not defined over the whole space. So if we attempt to define $(\exp t A)x$ by a power series $\sum_{k=0}^{\infty} \frac{(tA)^k}{k!}x$, we encounter some difficulties. First, we have to choose $x$ form $\bigcap_{k=0}^{\infty} \mathcal{D}(A^k)$ and we do not know how big this space is. Even if we do this, it will be difficult to prove the convergence of the series, let alone its convergence to $T_t x$. So we proceed to define the exponential in another way. As a preparation to the definition of the exponential function of an additive operator - not necessarily linear - we consider the exponential of a linear operator.

**Proposition.** Let $B$ be a linear operator from the Banach space $X$ into $X$. Then for each $x \in X$, $s - \lim_{n \to \infty} \sum_{k=0}^{n} \frac{B^k}{k!}x$ exists; denote this by $\exp Bx$. Then $\exp B$ is a linear operator and $\| \exp B \| \leq \exp(\| B \|)$.

**Proof.** We have $\| B^k \| \leq (\| B \|)^k$ $(k \geq 0)$. $\sum_{k=0}^{\infty} \frac{B^k}{k!}x$ is a Cauchy sequence; for $l > j$ we have

$$\left\| \sum_{k=0}^{l} \frac{B^k}{k!} - \sum_{k=0}^{j} \frac{B^k}{k!} \right\| = \left\| \sum_{k=j+1}^{l} \frac{B^k}{k!} \right\| \sum_{j+1}^{l} \frac{\| B \|^k}{k!}$$

and $\leq \sum_{k=0}^{\infty} \frac{\| B \|^k}{k!} \| x \|$ is convergent. So, by the completeness of the space, $s - \lim_{n \to \infty} \sum_{k=0}^{n} \frac{B^k}{k!}x$ exists; and the convergence is uniform in every sphere $\| x \| \leq M$; the above inequality shows that

$$\| \exp B \| \leq \exp(\| B \|) \| x \|.$$ 

$\square$
1. The exponential of a linear operator

So \( \exp B \) is a linear operator and

\[
\| \exp B \| \leq \exp(\| B \|).
\]

**Remark.** In a similar manner one can prove the following: Let a sequence of linear operators \( \{ S_n \} \), on a linear normed space \( X \) with values in a Banach space \( Y \) be a Cauchy sequence, i.e.,

\[
\lim_{n,m} \| S_n - S_m \| = 0.
\]

Then there exists a linear operator \( S \) forms \( X \) to \( Y \) such that \( \lim_{n \to \infty} \| S_n - S \| = 0 \) and \( \| S \| \leq \lim_{n \to \infty} \| S_n \| \).

**Theorem.** Let \( B \) and \( C \) be two linear operators from a Banach space \( X \) into \( X \). Assume that \( B \) and \( C \) commute, i.e., \( BC = CB \). Then

1) \( \exp B \cdot \exp C = E \cdot \exp (B + C) \)

2) \( D_t \exp(tB)x = s - \lim_{h \to \infty} \frac{\exp(t + h)B - \exp tB}{h} x \) exists and has the value \( B(\exp tBx) = (\exp tB).Bx \).

**Proof.** i) If \( \beta \) and \( \phi \) are complex numbers, we have

\[
\sum_{j=0}^{\infty} \frac{(t\beta)^j}{j!} \sum_{l=0}^{\infty} \frac{(t\phi)^l}{l!} = \frac{t(\beta + \phi)^m}{m!} \quad (t > 0);
\]

for, by the absolute convergence of each of the series on the left and the commutativity of \( \beta \) and \( \phi \) we may arrange the product on the left to be equal to the power series on the right. A similar proof holds when \( \beta \) and \( \phi \) are replaced by commuting linear operators \( B \) and \( C \) on a Banach space.

ii) Since \( tB \) and \( hB \) commute, we have by 1)

\[
\exp(t + h)B = \exp(tB) \cdot \exp(hB) = \exp(hB) \cdot \exp tB.
\]

So,

\[
\frac{\exp(t + h)B - \exp tB}{h} = \frac{\exp tB(\exp hB - I)}{h} = \frac{\exp(hB) - I}{h} \exp tB.
\]
iii) follows since
\[
\left\| \frac{\exp(hB) - I}{h} - B \right\| = \left\| \sum_{k=2}^{\infty} \frac{(hB)^k}{k!} \right\| \leq \sum_{k=2}^{\infty} \frac{B^k}{k!} h^{k-1} \to 0, \text{ as } h \to 0.
\]

\[\square\]

2 Representation of semi-groups

**Theorem.** Let A be the infinitesimal generator of a semi-group \( \{T_t\} \).

Then for each \( y \in X \)
\[
T_t y = s - \lim_{n \to \infty} \exp(tAJ_n)y
\]
uniformly in any bounded interval of t. (\( J_n \) is the resolvent \( (I - n^{-1}A)^{-1}, n > \beta \)).

**Proof.** \( (tAJ_n) = nt(J_n - I) \) is a linear operator and so \( \exp(tAJ_n) \) can be defined. Since \( ntI \) and \( ntJ_n \) commute we have
\[
(\exp tAJ_n) = \exp(-ntI). \exp(ntJ_n)
= \exp(-nt). \exp(ntJ_n).
\]

\[\square\]

Since \( \|J_n\| \leq 1/(1 - \beta n^{-1}) \) \( (n > \beta) \), we have
\[
\| \exp(tAJ_n) \| \leq \exp(-nt)\| \exp(ntJ_n) \|
\leq \exp(-nt) \exp(nt|J_n|)
\leq \exp(-nt) \exp(nt/1 - \beta n^{-1})
= \exp(tB/(1 - \beta n^{-1}))
\]

If \( x \in \mathcal{D}(A) \), \( D_T x = AT x = T x \) and hence
\[
D_t \{\exp[(t-s)AJ_n]\}T_s x = \exp((t-s)AJ_n)T_s Ax = \exp((t-s)AJ_n)AJ_n T_s x.
\]
Since $T_1T_s = T_sT_1 (= T_{r+s})$,

$$J_n = n \int_0^\infty e^{-nt}T_t dt$$

is the limit of Riemannian sums each of which commutes with each $T_s$; so $J_n$ commutes with each $T_s$ so that $AJ_n = n(J_n - I)$ commutes with each $T_s$. Now

$$T_t x - \exp(tAJ_n)x = [\exp((t - s)AJ_n)T_s x]_{s=0}^t$$

Since $\exp((t - s)AJ_n)T_s (A - AJ_n)x$ is strongly continuous in $s$, we have, for $x \in \mathcal{D}(A)$,

$$T_t x - \exp(tAJ_n)x = \int_0^t D_t \{ \exp((t - s)AJ_n)T_s x \} ds$$

$$= \int_0^t \exp((t - s)AJ_n)T_s (A - J_n Ax) ds$$

(as $AJ_n x = J_n Ax, axx \in \mathcal{D}(A)$)

So

$$\|T_t x - \exp(tAs_n)x\| \leq \int_0^t \| \cdots \| ds$$

$$\leq \int_0^t \| \exp(t-s)AJ_n \| \|T_s\| \|Ax - J_n Ax\| ds$$

$$\leq \|Ax - J_n Ax\| \int_0^t \exp \frac{\beta(t - s)}{1 - \beta n^{-1}} \exp \beta s ds$$

For each fixed $t_0 > 0$ and $n > \beta$, the integral is uniformly bounded for $0 \leq t \leq t_0$ as $n(> \beta) \to \infty$; also we know that for each $x \in X$, $s - \lim_{n \to \infty} J_n x = x$. Thus

$$T_t x = s - \lim_{n \to \infty} \exp(tAJ_n)x \text{ uniformly in } 0 \leq t \leq t_0,$$

if $x \in \mathcal{D}(A)$.

We now prove the formula for arbitrary $y \in X$. Since $\mathcal{D}(A)$ is dense in $X$, given $\varepsilon > 0$ we can find $x \in \mathcal{D}(A)$ such that $\|y - x\| \leq \varepsilon$. Then
\[ \|T_t y - \exp(tAJ_n y)\| \leq \|T_t y - T_t x\| + \|T_t x - \exp(tAJ_n x)\| \\
+ \|\exp(tAJ_n x) - \exp(tAJ_n y)\| \\
\leq \exp(\beta t) \varepsilon + \|T_t x - \exp(tAJ_n x)\| \\
+ \exp\left(\frac{t}{1 - n^{-1} \beta}\right) \varepsilon. \]

Since \( x \in \mathcal{D}(A) \), the middle term on the right tends to zero as \( n \to \infty \) uniformly in any bounded interval of \( t \). So

\[ \lim_{n \to \infty} \|T_t y - \exp(tAJ_n y)\| \leq 2 \exp(\beta t) \varepsilon, \]

and \( \varepsilon \) being arbitrary,

\[ T_t y = s - \lim_{n \to \infty} (\exp tAJ_n y), y \in X, \]

uniformly in any bounded interval of \( t \).

**Remark.** The above representation of the semi-group was obtained independently of \( E \). Hille who gave many representations in his book. One of them reads as follows:

\[ T_t x = s - \lim_{n \to \infty} \left( I - \frac{tA}{n}\right)^{-1} x \]

uniformly in any bounded interval of \( t \). It also shows the exponential character of the representation.
Lecture 8

1 An application of the representation theorem

In $C[0, \infty]$ consider $(T_t x)(s) = x(t + s)$. By the representation theorem

$$ (T_t x)(s) = x(t + s) = s - \lim_{n \to \infty} \exp (tA J_n x)(s) $$

$$ = s - \lim_{n \to \infty} \sum_{m=0}^{\infty} \frac{t^m}{m!} (A J_n)^m x(s) $$

uniformly in any bounded interval. From this we get an operational theoretical proof of the Weierstrass approximation theorem. Let $z(s)$ be a continuous function on the closed interval $[0, \alpha], 0 < \alpha < \infty$. Let $x(s) \in C[0, \infty]$ be such that $x(s) = z(s)$ for $s \in [0, \alpha]$ (such functions trivially exist). Put $s = 0$ in the above formula

$$ (T_t x)(0) = x(t) = s - \lim_{n \to \infty} \sum_{m=0}^{\infty} \frac{t^m}{m!} [(A J_n)^m x](0) $$

uniformly in $[0, \alpha]$. Thus shown that $z(s)$ is the uniform limit of polynomials on $[0, \alpha]$.

2 Characterization of the infinitesimal general of a semi-group

We next wish to characterize the infinitesimal generator of a semi-group by some of the properties we have established. First we prove the
**Proposition.** Let $A$ be an additive operator on a Banach space $X$ into itself with the following properties:

(a) $\mathcal{D}(A)$ is dense in $X$;

(b) there exists a $\beta \geq 0$ such that for $n > \beta$ the inverse $J_n = (I - n^{-1}A)^{-1}$ exists as a linear operator satisfying

$$
\|J_n\| \leq (1 - n^{-1}\beta)^{-1} (n > \beta).
$$

Then we have

i) $AJ_nx = n(J_n - I)x$, $x \in X$

ii) $AJ_nx = J_nAx = n(J_n - I)x$, $x \in \mathcal{D}(A)$

iii) $s - \lim_{n \to \infty} J_nx = x$, for $x \in X$.

*Proof.* i) and ii) are evident. To prove iii) let $y \in \mathcal{D}(A)$.

Then $y = J_ny - n^{-1}J_nAy$ and hence

$$
\|y - J_ny\| \leq n^{-1}\|J_n\|\|Ay\|
\leq n^{-1}(1 - n^{-1}\beta)^{-1}\|Ay\| \to 0 \text{ as } n \to \infty.
$$

Let $x \in X$. Since $\mathcal{D}(A)$ is dense in $X$, given $\varepsilon > 0$, there exists $y \in \mathcal{D}(A)$ such that $\|y - x\| \leq \varepsilon$. We then have

$$
\|x - J_nx\| \leq \|x - y\| + \|y - J_ny\| + \|J_ny - J_nx\|
\leq \varepsilon + \|y - J_ny\| + (1 - n^{-1}\beta)^{-1}\varepsilon.
$$

As $\|y - J_ny\| \to$ as $n \to \infty$,

$$
\lim_{n \to \infty}\|x - J_nx\| \leq \varepsilon,
$$

and $\varepsilon$ being arbitrary positive number, iii) is proved.
Theorem. An additive operator $A$ with domain $\mathcal{D}(A)$ dense in a Banach space $X$ and with values in $X$ is the infinitesimal generator of a uniquely determined semi-group $\{T_t\}$ with $\|T_t\| \leq e^{\beta t}$ if (and only if), for $n > \beta$, the inverse $J_n = (I - n^{-1}A)^{-1}$ exists as a linear operator satisfying $\|J_n\| \leq (1 - n^{-1}\beta)^{-1}$.

Proof. We put $T_t^{(n)} = (\exp tAJ_n)$. We have

$$\|T_t^{(n)}\| \leq \exp(-nt)\exp(nt\|J_n\|)$$

$$\leq \exp\frac{\beta t}{1 - n^{-1}\beta},$$

$$D_t T_t^{(n)} x = AJ_n T_t^{(n)} x = T_t^{(n)} AJ_n x, x \in X,$$

and

$$T_t^{(n)} x - x = \int_0^t T_s^{(n)} AJ_n x \, ds.$$ 

It is easy to see $J_n J_m = J_m J_n$; so $AJ_n = n(J_n - I)$ commutes with $T_t^{(m)} = \exp(tAJ_m)$. Thus, as in the proof of the representation theorem, we have, for any $x \in \mathcal{D}(A)$,

$$\|T_t^{(m)} x - T_t^{(n)} x\| = \left\| \int_0^t D_s\left\{ T_s^{(n)} - T_s^{(m)} \right\} x \, ds \right\|$$

$$= \left\| \int_0^t T_s^{(m)} T_s^{(m)} (AJ_m - AJ_n) x \, ds \right\| (\text{as } D_s T_s^{(m)} x = T_s^{(m)} AJ_m x)$$

$$\leq \| (J_m A - J_n A) x \| \int_0^t \exp \frac{\beta (t - s)}{1 - n^{-1}\beta} \exp \frac{\beta s}{1 - m^{-1}\beta} \, ds$$

So $\lim_{m,n \to \infty} \|T_t^{(m)} x - T_t^{(n)} x\| = 0$ uniformly in any finite interval of $t$. Let $y \in X$. Given $\varepsilon > 0$, there exists $x \in \mathcal{D}(A)$ such that $\|y - x\| \leq \varepsilon$. Then

$$\|T_t^{(m)} y - T_t^{(n)} y\| = \|T_t^{(m)} y - T_t^{(n)} x\| + \|T_t^{(m)} x - T_t^{(n)} x\|$$

$$+ \|T_t^{(n)} x - T_t^{(n)} y\|$$

$$\leq \exp\left(\frac{\beta t}{1 - m^{-1}\beta}\right) \varepsilon + \|T_t^{(m)} x - T_t^{(n)} x\| + \exp\left(\frac{\beta t}{1 - n^{-1}\beta}\right) \varepsilon.$$
So \( \lim_{m,n \to \infty} \|T_i^{(m)} y - T_i^{(n)} y\| \leq \epsilon. \) 2 \( \exp(\beta t) \) uniformly in any finite interval of \( t \). Therefore, by the completeness of \( X \), \( s - \lim_{n \to \infty} T_i^{(n)} y = T_i y \) exist and the convergence is uniform in any bounded interval of \( t \).

By the resonance theorem \( T_i \) is a linear operator; since \( T_i^{(n)} \) are strongly continuous in \( t \) and the convergence is uniform in any bounded interval of \( t \), \( T_i \) is strongly continuous in \( t \). Also,

\[
\|T_i\| \leq \lim_{n \to \infty} \|T_i^{(n)}\| \quad (\text{Cor. to response theorem })
\]

\[
\leq \exp(\beta t)
\]

We now prove that \( T_i T_s = T_{i+s}(T = I, \text{ evidently}) \).

Since \( T_i^{(n)} T_s^{(n)} = T_{i+s}^{(n)} \),

\[
\|T_{i+s} x - T_i T_s x\| \leq \|T_{i+s} x - T_i^{(n)} x\| + \|T_i^{(n)} x - T_i^{(n)} T_s x\| + \|T_i^{(n)} T_s x - T_i T_s x\|
\]

\[
\leq \|T_{i+s} x - T_i^{(n)} x\| + \exp \left( \frac{\beta t}{1 - n^{-1}\beta} \right) \|T_i^{(n)} x - T_i x\|
\]

\[
+ \|T_i^{(n)} T_s x - T_i T_s x\|
\]

\[
\to 0 \text{ as } n \to \infty.
\]

Finally let \( A' \) be the infinitesimal generator of the semi-group \( T_i \). We shall show that \( A' = A \). For this it is enough to prove that \( A' \) is an extension of \( A \) (i.e., \( x \in \mathcal{D}(A) \) implies \( x \in \mathcal{D}(A') \) and \( A' x = A x \)).

For, \((I - n^{-1}A')(n > \beta)\) maps \( \mathcal{D}(A') \) onto \( X \) in a one-one manner; by assumption \((I - n^{-1}A)\) maps \( \mathcal{D}(A') \) onto \( X \) in a one-one manner; but on \( \mathcal{D}(A), (I - n^{-1}A) = (I - n^{-1}A') \) and hence \( \mathcal{D}(A) = \mathcal{D}(A') \). To prove that \( A' \) is an extension of \( A \), we start with the formula

\[
T_i^{(n)} x - x = \int_0^t T_s^{(n)} A J_n x ds, \ x \in X.
\]

If \( x \in \mathcal{D}(A) \)

\[
\|T_s A x - T_s^{(n)} A J_n x\| \leq \|T_s A x - T_s^{(n)} A x\| + \|T_s^{(n)} A x - T_s A J_n x\|
\]

\[
\leq \|T_s - T_s^{(n)} A x\| + \exp \left( \frac{\beta s}{1 - n^{-1}} \right) \|A x - J_n A x\|
\]
(AJ_n x = J_n Ax, if x ∈ \mathcal{D}(A)).

As n → ∞ the first on the right tends to zero, uniformly in any bounded interval of \( s \); the second term also tends to zero, uniformly in any bounded interval of \( s \), as \( \exp \left( \frac{\beta s}{1 - \beta n^{-1}} \right) \) stays in such an interval and we know that

\[
s - \lim_{n \to \infty} J_n y = y, y \in X.
\]

Hence

\[
T_t x - x = \lim_{n \to \infty} (T_t^{(n)} x - x) = \lim_{n \to \infty} \int_0^t T_s^{(n)} A J_n x \, ds
\]

\[
= \int_0^t s - \lim_{n \to \infty} (T_s^{(n)} A J_n x) ds
\]

\[
= \int_0^t T_s A x ds
\]

(using the uniformly of convergence in \([\alpha, t]\)). Therefore

\[
s - \lim_{n \to \infty} \frac{T_t x - x}{t} = T_0 A x = A x.
\]

i.e., if x ∈ \mathcal{D}(A) then x ∈ \mathcal{D}(A') and A'x = Ax.

The uniqueness of the semi-group \( \{T_t\} \) with A as the infinitesimal generator follows from the representation theorem for semi-groups proved earlier.
Lecture 9

1 Group of operators

We add certain remarks which will be useful for the application of semi-group theory to Cauchy’s problem. The first of these relates to conditions under which a semi-group becomes a group; this will be useful in connection with the wave equation.

**Definition.** A one parameter family \(T_t\) of linear operators \(T_t\) of a Banach space \(X\) is called a group of linear operators of normal type (or simply a group) if the following conditions are satisfies:

i) \(T_tT_s = T_{t+s}\), \(T_o = I\) (group property)

ii) \(s \to \lim_{n \to 0} T_{t/n}x = T_{t_0}x\) for each \(x \in X\) and \(t_0 \in (-\infty, \infty)\)

iii) there exists a \(\beta \geq 0\) such that for all \(t\)

\[ ||T_t|| \leq e^{\beta|t|}.\]

(The infinitesimal generator of a group is defined by: \(Ax = \lim_{t\to 0} \frac{T_t x - x}{t}\)).

**Theorem.** Let \(A\) be an additive operator from a Banach space \(X\) into \(X\) such that \(\mathcal{D}(A)\) is dense in \(X\). A necessary and sufficient condition that \(A\) be the infinitesimal generator of a group \(T_t\) is that there exists a \(\beta \geq 0\) such that for every \(n\) with \(|n| > \beta\) the inverse \(J_n = (I - N^{-1}A)^{-1}\) exists as linear operator with \(||J_n|| \leq \beta/(1 - |n|^{-1}\beta)\).
Proof. Necessity. Let \( \{T_t\} \) be a group. Consider the two semi-groups \( \{T_t\}_{t \geq 0} \), \( \{\hat{T}_t\}_{t \geq 0} \) where \( \hat{T}_t = T_{-t} \). The infinitesimal generator of the semi-group \( \{T_t\}_{t \geq 0} \) coincides with the infinitesimal generator \( A \) of the group; let \( A' \) be the infinitesimal generator of \( \{\hat{T}_t\} \). □

If we show that \( A' = -A \) the proof of the necessity part will be complete. Let \( x \in \mathcal{D}(A') \). Then

\[
\frac{\hat{T}_h - I}{h} x = A' x.
\]

Putting \( x_n = h^{-1}(\hat{T}_h - I)x \), we have

\[
||T_h x_h - A' x|| \leq ||T_h x_h - T_h A' x|| + ||T_h A' x - A' x|| \leq ||T_h|| ||x_h - A' x|| + ||T_h A' x - A' x|| \leq (\exp \beta h) ||x_h - A' x|| + ||T_h A' x - A' x|| \rightarrow 0 \text{ as } h \downarrow 0.
\]

Thus for \( x \in \mathcal{D}(A') \)

\[
-Ax = s - \lim_{n \uparrow 0} h^{-1}(I - T_h) = s - \lim_{n \uparrow 0} T_h x_h = A' x.
\]

Hence \( x \in \mathcal{D}(A') \) implies \( x \in \mathcal{D}(A) \) and \( A' x = -Ax \). Similarly it is proved that if \( x \in \mathcal{D}(A) \), then \( x \in \mathcal{D}(A') \) and \( A' x = -Ax \). So \( A' = -A \).

Sufficiency: We can construct two semi-groups \( \{T_t\}_{t \geq 0} \) and \( \{\hat{T}_t\}_{t \geq 0} \) as follows:

\[
T_t x = s - \lim_{n \rightarrow \infty} T_t^{(n)} x = s - \lim_{n \rightarrow \infty} \exp(tAJ_n)x = -s - \lim_{n \rightarrow \infty} \exp(n(t - n^{-1}A)^{-1} - I)x
\]

\[
\hat{T}_t x = s - \lim_{n \rightarrow \infty} \exp(t - AJ_n)x = s - \lim_{n \rightarrow \infty} \exp(n(t + n^{-1}A)^{-1} - I)x
\]

If we show that \( \hat{T}_t T_t = T_t \hat{T}_t = I \), then

\[
\hat{T}_t = \begin{cases} 
T_t & \text{for } t \geq 0 \\
\hat{T}_{-t} & \text{for } t \leq 0 
\end{cases} \quad (-\infty < t < \infty)
\]
will be a group with $A$ as the infinitesimal generator. Since $J_n = (I - n^{-1}A)^{-1}$ commutes with $J_{-n} = (I + n^{-1}A)^{-1}$ we have
\[
(I - n^{-1}A)^{-1} + (I + n^{-1}A)^{-1} = [(I + n^{-1}A) + (I - n^{-1}A)](I - n^{-1}A)^{-1}(I + n^{-1}A)^{-1} = 2(I - n^{-1}A)^{-1}(I + n^{-1}A)^{-1} = 2(I - n^{-2}A^2)^{-1}.
\]

Since $J_k$ maps $X$ onto the dense subspace $\mathcal{D}(A)$ of $X$, $J_n J_{-n} = (I - n^{-1}A^2)^{-1}$ maps $X$ onto a dense subspace $\mathcal{D}(A^2)$. Moreover
\[
\| (I - n^{-2}A)^{-1} \| \leq \| J_n \| \| J_{-n} \| \leq (1 - \beta/n)^{-1} \left( 1 - \frac{\beta}{n} \right)^{-1} = (1 - \beta^2/n^2)^{-1}.
\]

Therefore $A^2$ is the infinitesimal generator of a semi-group $\exp(tA^2)$.
\[
\exp(tA^2)x = s - \lim_{m \to \infty} \exp(tA^2(I - m^{-1}A^2)^{-1})x = s - \lim_{m \to \infty} \exp(m^2t[(I - m^{-1}A^2)^{-1} - I])x
\]
the convergence being uniform in $t$ in any finite interval of $t$.

We have
\[
\| T_{u_{n,k}}^{-1} T_{u_{n,k}} x - T_{u_{n,k}}^{-1} T_{u_{n,k}} x \| \leq \| T_{u_{n,k}}^{-1} T_{u_{n,k}} x - T_{u_{n,k}}^{-1} T_{u_{n,k}} x \| + \| T_{u_{n,k}}^{-1} T_{u_{n,k}} x - T_{u_{n,k}}^{-1} T_{u_{n,k}} x \|
\]
\[
\leq \left\| T_{u_{n,k}}^{-1} T_{u_{n,k}} x \right\| + \left\| T_{u_{n,k}}^{-1} T_{u_{n,k}} x \right\|
\]
\[
\to 0 \text{ as } n \to \infty,
\]
uniformly in $t$ in any bounded interval of $t$.

That the first on the right tends to zero uniformly in $t$ in any bounded interval of $t$ may be proved as follows: Let $0 \leq t \leq t_0 < \infty$. For any $\varepsilon > 0$, we can find $t_1, \ldots, t_k, 0 \leq t_1, \ldots, t_k \leq t$, such that
\[
\inf_{1 \leq s \leq k} \| T_{u_{n,k}}^{-1} T_{u_{n,k}} x \| \leq \varepsilon,
\]
(by the strong continuity of $T_t$ in $t$).
Now
\[ \|(T_i - T_i^{(n)})\hat{T}_i x\| \to 0 \quad (i = 1, 2, \ldots, k) \]
uniformly in \( t \) for \( 0 \leq t \leq t_o \), and hence, choosing \( t_i \) properly for given \( t \), we have
\[
\|(T_i - T_i^{(n)})\hat{T}_i x\| \leq \|(T_i - T_i^{(n)})\hat{T}_i x\| + \exp \beta t + \exp \frac{\beta t}{1 - n^{-1} - \beta} \varepsilon.
\]
So the right side tends to zero uniformly in \( 0 \leq t \leq t_o \).

Since
\[
T_i^{(n)}\hat{T}_i x = \exp \left( n \left( (I - n^{-1} A)^{-1} + (I + n^{-1} A)^{-1} - 2I \right) \right) x
\]
\[
= \exp \left( \frac{2t}{n} n^2 \left( (I - n^{-2} A^2)^{-1} - I \right) \right) x,
\]
we have
\[
T_i\hat{T}_i x = s - \lim_{n \to \infty} \exp \left( \frac{2t}{n} n^2 \left( (I - n^{-2} A^2)^{-1} - I \right) \right) x
\]
the convergence being uniform in any bounded interval of \( t \). Thus
\[
T_i\hat{T}_i x = \exp(0. A^2 x) = x.
\]
Similarly
\[
\hat{T}_i T_i x = x.
\]


Lecture 10

1 Supplementary results

We shall now prove some results which supplement our earlier results; these will be useful in applications.

**Theorem.** 1. For a semi-group \( \{T_t\} \) the infinitesimal generator \( A \) may be defined by

\[
\lim_{h \to 0} \frac{T_h - I}{h} x.
\]

i.e., if \( \tilde{A} \) is the operator with \( \mathcal{D}(\tilde{A}) = \{ x \mid \lim_{h \to 0} \frac{T_h - I}{h} x \text{ exists} \} \) and

\[
\tilde{A}x = \lim_{h \to 0} \frac{T_h - I}{h} x,
\]

then \( \tilde{A} = A \).

2. If \( \{T_t\}_{t \geq 0} \) is a family of linear operators on a Banach space \( X \) such that \( T_{t+s} = T_t T_s, T_0 = I \) and \( \|T_t\| \leq e^{\beta t}, \beta \geq 0 \) then the following two conditions are equivalent:

(i) strong continuity of \( T_t \), i.e., \( \lim_{t \to t_0} \frac{T_t - T_0}{t - t_0} x = T_{t_0} x \) for each \( t_0 \geq 0 \) and \( x \in X \).

(ii) weak right continuity at \( t = 0 \), i.e., \( \lim_{h \to 0} \frac{T_t - I}{h} x = x \), for \( x \in X \).

3. The infinitesimal generator is a semi-group is a closed operator.

**Proof.** It is evident that \( \tilde{A} \) is an extension of \( A \). We shall show that \( A \) is an extension of \( \tilde{A} \), i.e., if \( x \in \mathcal{D}(\tilde{A}) \), then \( x \in \mathcal{D}(A) \) and \( Ax = \tilde{A}x \). If
(For, if \( w - \lim_{h \downarrow 0} T x h = y \), and \( T \) is a linear operator, then \( w - \lim_{h \downarrow 0} T x h = T y \); in fact, if \( f \in X^* \), \( \hat{f}(y) = f(T Y) \) is a linear functional on \( X \), as \( |\hat{f}(y)| \leq ||f|| ||T_y|| \leq ||f|| ||T|| ||y|| \), and \( f(T Y) - f(T x h) = \hat{f} y - \hat{f} x h \to 0 \) as \( h \downarrow 0 \). So, if \( x \in \mathcal{D}(A) \), \( f(T x) \) has right derivative \( \frac{df}{dt}(T x) = f(T \tilde{A} x) (t \geq 0) \), which is continuous for \( t \geq 0 \), by the strong continuity of \( T_t \). Therefore the derivative \( \frac{df}{dt}(T x) \) exists for each \( t \geq 0 \) and is continuous. 

So

\[
\begin{align*}
 f(T x - x) &= f(T x) - f(x) = \int_0^t f(T \tilde{A} x) ds \\
 &= f\left( \int_0^t T \tilde{A} x ds \right), \text{ for each } f \in X^*.
\end{align*}
\]

Continuously, by the Hahn-Banach theorem,

\[
T x - x = \int_0^t T \tilde{A} x ds.
\]

Since \( T_t \) is strongly continuous in \( t \) it follows that

\[
\begin{align*}
 s - \lim_{t \downarrow 0} \frac{T_t - I}{t} x = T_0 \tilde{A} x = \tilde{A} x.
\end{align*}
\]

Thus if \( x \in \mathcal{D}(\tilde{A}) \), then \( x \in \mathcal{D}(A) \) and \( \tilde{A} x = Ax \).

**PROOF.** Evidently (i) implies (ii). To prove that (ii) implies (i), let \( x_0 \) be a fixed element of \( X \). We shall show that \( w - \lim_{t \downarrow t_0} T x_0 = T_0 x_0 \) for each \( t \geq 0 \). Consider the function \( x(t) = T_t x_0 \). For \( t_0 \geq 0 \), \( x(t) \) is right continuous at \( t_0 \), as \( w - \lim_{t \downarrow t_0} T x_0 = w - \lim_{h \downarrow 0} T h T_0 x_0 \). \( x(t) \) has the following three properties:
1. Supplementary results

(a) $x(t)$ is weakly measurable, i.e., for any $f \in X^*$, $f(x(t))$ is measurable (since a right continuous numerical function is measurable).

(b) $\|x(t)\|$ is bounded in any bounded interval of $t$.

(c) there exists a countable set $M = \{x_n\}$ such that $x(t) \geq 0$ is contained in the closure of $M$.

To prove (c), let $\{t_k\}$ be the totality of positive rational numbers. Consider finite linear combinations $\sum \alpha_k x(t_k)$ where $\alpha_k$ are rational numbers if $X$ is real and if $X$ is complex $\alpha_k = a_k + ib_k$ with $a_k$ and $b_k$ rational. These elements form a countable set $M = \{x_n\}$. The closure of $M, \bar{M}$, contains $x(t)$, for each $t \geq 0$.

For, if not, let $t_0 \geq 0$ be a number such that $x(t_0)$ does not belong to $\bar{M}$. $\bar{M}$ is a closed linear subspace of $X$. By the Hahn-Banach theorem, there exists a linear functional $f_\circ$ on $X$ such that $f_\circ(x(t_0)) \neq 0$ and $f_\circ(x) = 0$ for $x' \in \bar{M}$. Take a sequence $t'_k \downarrow t_0$ ($t'_k$ positive rational). By the weak right continuity of $x(t)$ at $t_0$,

$$f_\circ(x(t'_k)) \rightarrow f_\circ(x(t_0)).$$

But $f_\circ(x(t'_k)) = 0$ and $f_\circ(x(t_0)) \neq 0$. We have thus arrived at a contradiction.

We next prove a result, due to N. Dunford (On one parameter group of linear transformations, Ann. of Math., 39(1938), 569 – 573), according of which the properties (a), (b) and (c) listed above imply the strong continuity of $x(t)$. First we show that $\|x(t)\|$ is measurable in $t$. Let $f_n \in X^*$ be such that $f_n(x_n) = \|x_n\|$ and $\|f_n\| = 1$. Let $f(t) = \sup_{n \geq 1} f_n(x(t))$; since each $f_n(x(t))$ is measurable, $f(t)$ is measurable in $t$. But $\|x(t)\| = f(t)$; for

$$f(t) \geq |f_n(x(t))| \geq |f_n(x_n)| - |f_n(x(t) - x_n)|$$

$$\geq \|x_n\| - \|x(t) - x_n\|$$

and $x(t)$ is in the closure of the set $M$ so that $f(t) \geq \|x(t)\|$; since $|f_n(x(t))| \leq \|x(t)\|$, $f(t) \leq \|x(t)\|$. Thus $f(t) = \|x(t)\|$ and $\|x(t)\|$ is measurable.
By a similar argument, \( \|x(t) - x_n\| \) is measurable in \( t \) for each \( n \). It follows, using (c), that the half-line \([0 \leq t < \infty)\) can be represented, for each integer \( m \), as a countable union of measurable sets \( S_{m,n} \).

\[
[0, \infty) = \bigcup_{n=1}^{\infty} S_{m,n}, S_{m,n} = \{ t : \|x(t) - x_n\| \leq m^{-1} \}
\]

If we define

\[
S'_{m,1} = S_{m,1}, \ldots, S'_{m,n} = S_{m,n} - \bigcup_{k=1}^{n-1} S'_{m,k},
\]

we have a decomposition of \([0, \infty)\) into disjoint measurable sets \( S'_{m,n} (n = 1, 2, \ldots) \) such that \( \|x(t) - x_n\| \leq m^{-1} \) in \( S'_{m,n} \).

Therefore the strongly measurable step-function (i.e., a countably valued function taking each of its values exactly on a measurable set)

\[ x^m(t) = x_n \text{ for } t \in S'_{m,n} \]

converges to \( x(t) \) as \( m \to \infty \) uniformly in \([0, t)\). Thus \( x(t) \) is a strongly measurable function, a strongly measurable function being a functional which is the uniform limit of a sequence of strongly measurable step functions. We may then define the Bochner integral of \( x(t) \) by:

\[
\int_{\alpha}^{\beta} x(t) dt = s - \lim_{\rightarrow \infty} \int_{\alpha}^{\beta} x^{(m)}(t) dt, \quad 0 \leq \alpha < \beta < \infty
\]

(\( \int_{\alpha}^{\beta} x^m(t) dt \) may be defined, as in the case of the ordinary Lebesgue integral, as the strong limit of finitely valued functions, each taking each of its values exactly on a measurable set). We have

\[
\| \int_{\alpha}^{\beta} x(t) dt \| \leq \int_{\alpha}^{\beta} \|x(t)\| dt.
\]

Let \( 0 \leq \alpha < \eta < \beta < \xi - \varepsilon < \xi (\varepsilon > 0) \).
1. Supplementary results

Since
\[ x(\xi) = T_\xi x_0 = T_\eta T_{\xi-\eta} x_0 = T_\eta x(\xi - \eta), \]
we have
\[ (\beta - \alpha) x(\xi) = \int_\alpha^\beta x(\xi) d\eta = \int_\alpha^\beta T_\eta (\xi - \eta) d\eta, \]
the integrals being Bochner integrals. So
\[ (\beta - \alpha) \{ x(\xi \pm \varepsilon) - x(\xi) \} = \int_\alpha^\beta T_\eta \{ x(\xi \pm \varepsilon - \eta) - x(\xi - \eta) \} d\eta. \]

Thus
\[ |\beta - \alpha| ||x(\xi \pm \varepsilon) - x(\xi)|| \leq \sup_{\alpha \leq \eta \leq \beta} ||T_\eta|| \int_\xi^\beta ||x(\tau \pm \varepsilon) - x(\tau)|| d\tau. \]

But the right side tends to zero as \( \varepsilon \downarrow 0 \). (This we see by approximating \( x(\xi) \), in bounded interval, uniformly with bounded, finitely valued strongly measurable functions. For, then the result is reduced to the case of numerical measurable step functions.) Thus \( x(\xi) \) is strongly continuous for \( \xi > 0 \).

To prove the strong continuity at \( \xi = 0 \) we proceed as follows: For positive rational \( t_k \), since
\[ T_\xi x(t_k) = T_{\xi+t_k} x_0 = T_{\xi+t_k} x_0 = x(\xi + t_k), \]
we have, using the continuity for \( \xi > 0 \) proved above,
\[ s - \lim_{\xi \uparrow 0} T_\xi x(t_k) = x(t_k). \]

It follows that \( s - \lim_{\xi \uparrow 0} T_\xi x_n = x_n \) for each \( x_n \); also \( x(t), t \geq 0 \), in particular \( x(0) = x_0 \), belongs to \( \tilde{M} (M = \{ x_n \}) \). It follows therefore, from the inequalities,
\[
||x(\xi) - x_0|| \leq ||T_\xi x_n - x_n|| + ||x_n - x_0|| + ||T_\xi (x_0 - x_n)|| \\
\leq ||T_\xi x_n - x_n|| + ||x_n - x_0|| + \sup_{0 \leq \xi \leq 1} ||T_\xi|| ||x_0 - x_n||,
\]
that \( \lim_{\xi \downarrow 0} x(\xi) = x_0 \), i.e., \( T_\xi \) is strongly continuous at \( \xi = 0 \).
**PROOF.** An additive operator $A$ (with domain $\mathcal{D}(A)$) is said to be closed if it possesses the following property: if $\{x_n\}$ is a sequence of elements of $\mathcal{D}(A)$ such that $s - \lim_{n \to \infty} x_n = x$ and $s - \lim_{n \to \infty} Ax_n = y$, then $x$ belongs to $\mathcal{D}(A)$ and $Ax = y$. Evidently a linear operator is closed.

To prove (3) let $k > \beta$. Then $J_k = \left(I - \frac{A}{k}\right)^{-1}$ is a linear operator. Let $\{x_n\}$ be a sequence, $x_n \in \mathcal{D}(A)$ such that $s - \lim_{n \to \infty} x_n = x$, $s - \lim_{n \to \infty} Ax_n = y$. Then $s - \lim_{n \to \infty} \left(x_n - \frac{A}{k} x_n\right) = x - \frac{y}{k}$. By the continuity of $J_k$, $s - \lim_{n \to \infty} J_k \left(x_n - \frac{A}{k} x_n\right) = J_k \left(x - \frac{y}{k}\right)$, i.e., $x = J_k \left(x - \frac{y}{k}\right)$. So $x \in \mathcal{D}(A)$. Since

$$\left(I - \frac{A}{k}\right)x = \left(I - \frac{A}{k}\right)J_k \left(x - \frac{y}{k}\right) = x - \frac{y}{k},$$

we have $Ax = y$.

**Remark.** It is to be noted that the theory has been extended for $\{T_t\}_{t > 0}$ satisfying

$$T_t T_s = T_{t+s}$$

and the strong continuity in $t$ for $t > 0$. 

Lecture 11

1 Temporally homogeneous Markoff process on a locally compact topological space

Let $R$ be a locally compact topological space, countable at infinity. We consider in $R$ 'a probabilistic movement'. Suppose that for each triple $(t, x, E)$ consisting of a real number $t > 0$, a point $x \in R$ and Borel set $E \subset R$ there is given a real number $P(t, x, E)$ such that the following conditions are satisfied.

i) $P(t, x, E) \geq 0$, $P(t, x, R) = 1$

ii) for fixed $t$ and $x$, $P(t, x, E)$ is a countably additive set function on the Borel sets

iii) for fixed $t$ and $E$, $P(t, x, E)$ is a Borel measurable function in $x$

iv) $P(t + s, x, E) = \int_R P(t, x, dy) P(s, y, E) t, s > 0$. (Chapman - Kolmogoroff relation).

The function $P(t, x, E)$ is called the transition probability; this gives the probability that, in this process, a point $x \in R$ is transferred to the Borel set $E$ after $t$ units of time. We say then that there is given a temporally homogeneous Markoff process on $R$ (temporal homogeneity means that the motion does not depend on the initial time but only on the time elapsed).
2 Brownian motion on a homogeneous Riemannian space

Next, we wish to define the ‘spatial homogeneity’ of the process. We assume that $R$ is an $n$-dimensional, orientable connected $C^\infty$ Riemannian space such that the (full) group of isometries $G$ of $R$, which is a Lie group, is transitive on $R$ (i.e., for each pair $x, y \in R$ there exists an isometry $S^x$ such that $S^x x = y$. The process $P(t, x, E)$ is called spatially homogeneous if

\[ P(t, x, E) = P(t, S^x x, S^x E) \text{ for each } S^x \in G, x \in R, E \subset R. \]

A temporally and spatially homogeneous Markov process on $R$ is called a Brownian motion on $R$, if the following condition, known as the continuity condition of Lindeberg, is satisfied.

\[ \lim_{t \downarrow 0} t^{-1} \int_{\text{dist}(x, y) > \varepsilon} P(t, x, dy) = 0, \text{ for every } \varepsilon > 0 \text{ and } x \in R. \]

Proposition. Let $C[R]$ denote the Banach space of bounded uniformly continuous real valued functions $f(x)$ on $R$, with the norm

\[ \|f\| = \sup_{x \in R} |f(x)|. \]

Define

\[ (T_t f)(x) = \begin{cases} \int_{R} P(t, x, dy) f(y), & \text{if } t > 0, \\ f(x), & \text{if } t = 0. \end{cases} \]

Then $T_t$ defines a semi-group of normal type in $C[R]$.

Proof. We have by condition $(i)$,

\[ |T_t f(x)| \leq \sup_{y \in R} |f(y)|. \]

\[ \square \]
If we define a linear operator $S$ by $(S f)(x) = f(S^* x), S^* \in G$, we have $T S = S T$. For,

$$(ST_t f)(x) = (T f)(S^* x)$$

$$= \int P(t, S^* x, dy) f(y)$$

$$= \int P(t, S^* x, d(S^* y)) f(S^* y)$$

$$= \int P(t, x, dy) f(S^* y) = (TS f)(x).$$

If $S^* \in G$ be such that $S^* x = x^1$, we have

$$(T_t f)(x) - (T_t f)(x') = (T_t f)(x) - (ST_t f)(x)$$

$$= T_t (f - S f)(x).$$

By the uniform of continuity of $f(x)$ and the above equality, we see that $(T_t f)(x)$ is uniformly continuous and bounded. The semi-group property follows easily from Fubini’s theorem and the Chapman-Kolmogorff relation $(T_0 = I$ by definition).

To prove the strong continuity, it is enough by and earlier theorem, to verify weak right continuity at $t = 0$. Since the conjugate space of $C[R]$ is the space of measures of finite total variation, it is enough to show that $\lim_{t \downarrow 0} (T_t f(x)) = f(x)$ boundedly in $x$.

Now

$$|(T_t f)(x) - f(x)| = \left| \int \int P(t, x, dy)[f(y) - f(x)] \right| \text{ by(i)}$$

$$= \left| \int_{R} P(t, x, dy)[f(y) - f(x)] \right| + \left| \int P(t, x, dy)[f(y) - f(x)] \right|$$

$$\leq \left| \cdots \cdots \cdots \cdots \right| + 2\|f\| \int_{dist(x,y) > e} P(t, x, dy)$$

$$\leq 1.$$
The first term on the right tends to zero as \( \varepsilon \to 0 \) and, for fixed \( \varepsilon \), the second term tends to zero boundedly in \( x \) as \( t \downarrow 0 \) (by \((vi)\)), and the spatial homogeneity). Thus \( \lim_{t \downarrow 0} (T_t f)(x) = f(x) \) boundedly in \( x \).

**Theorem.** Let \( x_o \) be a fixed point of \( R \). Let us assume that the isotropy group \( G_o = \{ S^* | S^* \in G, S^* x_o = x_o \} \) is compact. (\( G_o \), being a closed subgroup of Lie group, is a Lie group). Let \( A \) be the infinitesimal generator of \( T_t \). Then

(i) if \( f \in \mathcal{D}(A) \cap C^2 \) (\( C^2 \) denoting the set of twice continuously differentiable functions), then, for a coordinate system \((x^1 \cdots x^n)\) at \( x_o \),

\[
(Af)(x_o) = a^i(x_o) \frac{\partial f}{\partial x^i_o} + b^{ij}(x_o) \frac{\partial^2 f}{\partial x^i_o \partial x^j_o}
\]

(adapting the summation convention), where

\[
a^i(x_o) = \lim_{t \downarrow 0} t^{-1} \int \limits_{\text{dist}(x_o,x) \leq \varepsilon} (x^i - x^i_o)P(t, x_o, dx)
\]

\[
b^{ij}(x_o) = \lim_{t \downarrow 0} t^{-1} \int \limits_{\text{dist}(x_o,x) \leq \varepsilon} (x^i - x^i_o)(x^j - x^j_o)P(t, x_o, dx)
\]

the limits existing independently of sufficiently small \( \varepsilon > 0 \).

(ii) The set \( \mathcal{D}(A) \cap C^2 \) is 'big' in the sense that, for any \( C^2 \) function with compact support there exists \( f(x) \in \mathcal{D}(A) \cap C^2 \) such that

\[
f(x_o), \frac{\partial f}{\partial x^i_o}, \frac{\partial^2 f}{\partial x^i_o \partial x^j_o}
\]

\[
g(x_o), \frac{\partial g}{\partial x^i_o}, \frac{\partial^2 g}{\partial x^i_o \partial x^j_o}
\]

are arbitrarily near respectively.

**Proof.**

**Step 1.** Let \( g(x) \) be a \( C^\infty \) function with compact support.

If \( f \in \mathcal{D}(A) \), the convolution

\[
(f \otimes g)(x) = \int_G f(S^*_y x)g(S^*_y x)dy,
\]
(\(S^*_y\) denotes a generic element of \(G\) and \(dy\) a fixed right invariant Haar measure on \(G\)) is \(C^\infty\) and belongs to \(\mathcal{D}(A)\). (The integral exists since the isotropy group is compact and \(g\) has compact support). By the uniform continuity of \(f\) and the compactness of the support of \(g\) we can approximate the integral by Riemann sums \(\sum_{i=1}^k f(S^*_y x)C_i\) uniformly in \(x\): 

\[
(f \otimes g)(x) = s - \lim_{n \to \infty} \sum_{i=1}^k f(S^*_y x_i)C_i.
\]

Since \(T_i S = S T_i\), \(S\) commutes with \(A\), i.e., if \(f \in \mathcal{D}(A)\), then \(S f \in \mathcal{D}(A)\) and \(AS f = S A f\). Putting \(h(x) = (A f)(x)\), \((h \in C[R])\).

\[
A \left( \sum_{i=1}^m f(S^*_y x)C_i \right) = \sum_{i=1}^m (A S y_i f)(x)C_i = \sum_{i=1}^m (S y_i A f)(x)C_i = \sum_{i=1}^m h(S^*_y x)C_i
\]

and the right hand side tends to \((h \otimes g)(x) = (A f \otimes g)(x)\). Since \(A\) is closed, it follows that \(f \otimes g \in \mathcal{D}(A)\), and \(A(f \otimes g) = A f \otimes g\). Since \(R\) is a homogeneous space of the Lie group \(G\) (by the closed subgroup \(G_o\)) we can find a coordinate neighbourhood \(U\) of \(x_o\) and for each \(x \in U\) an element \(S^*(x) \in G\) such that \(i) S^*(x) = x_o\) ii) \(S^*(x)x_o\) depends analytically on the coordinate functions \(x^1 \ldots x^n\). by the right invariance of the Haar measure,

\[
(f \otimes g)(x) = \int_G f(S^*_y S^*(x)x_o)g(S^*_y S^*(x)x_o)dy = \int_G f(S^*_y x_o)g(S^*_y S^*(x)x_o)dy, \quad x \in U.
\]

The function on the right side is \(C^\infty\) in a neighbourhood of \(x_o\) and

\[
\frac{\partial^{q_1 \ldots q_n}}{\partial(x^1)^{q_1} \cdots (x^n)^{q_n}} f \otimes g(x) = \int_G f(S^*_y x_o) \frac{\partial^{p_1 \ldots p_n}}{\partial(x^1)^{p_1} \cdots (x^n)^{p_n}} g(S^*_y S^*(x)x_o)dy
\]

\(\square\)
Lecture 12

1 Brownian motion on a homogeneous Riemannian space (Contd.)

Proof.

Step 2. Remarking that $\mathcal{D}(A)$ is dense in $C[R]$ and choosing $f$ and $g$ properly we obtain

(a) there exist $C^\infty$ functions $F^1(x), \ldots, F^n(x) \in \mathcal{D}(A)$ such that the Jacobian
\[ \frac{\partial(F^1(x), \ldots, F^n(x))}{\partial(x^1, \ldots, x^n)} > 0 \text{ at } x_o. \]

(b) there exists a $C^\infty$ function $F_o(x) \in \mathcal{D}(A)$ such that
\[ (x^i - x_{i,o})(x^j - x_{j,o}) \frac{\partial^2 F}{\partial x^i_o \partial x^j_o} \geq \sum_{i=1}^{n} (x^i - x_{i,o})^2. \]

We can use $F^1(x), \ldots, F^n(x)$ as coordinate functions in a neighbourhood $d(x_o, x) < \varepsilon$; we denote these new local coordinates by $(x_1, \ldots, x_n)$.

Since $F^j(x) \in \mathcal{D}(A)$,
\[ \lim_{t \downarrow t_o} \frac{T_t F^j(x) - F^j(x)}{t} \]
exists and $= AF^j(x)$
\[ (AF^j)(x) = \lim_{t \downarrow t_o} t^{-1} \int_R P(t, x_o, dx)(F^j(x) - F^j(x_o)) \]
\[\lim_{t \to 0} t^{-1} \int_{d(x, x_0) \leq \varepsilon} P(t, x, dx) (F(x) - F(x_0))\]

independent of \(\varepsilon > 0\), by Lindeberg’s condition. So, for the coordinate functions \(x^1 \cdots x^n, (x^i = F^i)\),

\[\lim_{t \to 0} t^{-1} \int_{d(x, x_0) \leq \varepsilon} (x^i - x_0^i) P(t, x_0, dx) = a^i(x_0)\]

independent of \(\varepsilon > 0\). Since \(F_0 \in \mathcal{A}(A)\), we have, using Lindeberg’s condition,

\[(AF_0)(x_0) = \lim_{t \to 0} t^{-1} \int_{R} P(t, x_0, dx)(F(x) - F_0(x_0))\]

\[= \lim_{t \to 0} \int_{d(x, x_0) \leq \varepsilon} P(t, x_0, dx)(F(x) - F_0(x_0))\]

\[= \lim_{t \to 0} \left[ t^{-1} \int_{d(x, x_0) \leq \varepsilon} (x^i - x_0^i) \frac{\partial F_0}{\partial x_0^i} P(t, x_0, dx) \right.\]

\[+ t^{-1} \int_{d(x - x_0) \leq \varepsilon} (x^i - x_0^i)(x^j - x_0^j) \left( \frac{\partial^2 F_0}{\partial x^i \partial x^j} \right) P(t, x_0, dx)\]

\[x = x_0 + \Theta(x - x_0, 0 < \Theta 1.\]

The first term on the right has a limit \(a^i(x_0) \frac{\partial F_0}{\partial x_0^i}\); hence by the positivity of \(P\), and \((b)\),

\[\lim_{t \to 0} t^{-1} \int_{d(x, x_0) \leq \varepsilon} \sum_{i=1}^{n} (x^i - x_0^i)^2 P(t, x_0, dx) < \infty \quad (*)\]

**Step 3.** Let \(f \in \mathcal{A}(A) \cap C^2\). Then, expanding \(f(x) - f(x_0),\)

\[\frac{T_t f(x_0) - f(x_0)}{t} = t^{-1} \int_{R} f(x) - f(x_0) P(t, x_0, dx)\]
\[ t^{-1} \int_{d(x,x_o) \geq \varepsilon} f(x) - f(x_o) P(t, x_o, dx) \]
\[ + t^{-1} \int_{d(x,x_o) \leq \varepsilon} (x^i - x_o^i) \frac{\partial f}{\partial x^i_o} P(t, x_o, dx) \]
\[ + t^{-1} \int_{d(x,x_o) \leq \varepsilon} (x^i - x_o^i)(x^j - x_o^j) \frac{\partial^2 f}{\partial x^i_o \partial x^j_o} P(t, x_o, dx) \]
\[ + t^{-1} \int_{d(x,x_o) \leq \varepsilon} (x^i - x_o^i)(x^j - x_o^j) C_{ij}(\varepsilon) P(t, x_o, dx) \]
\[ = C_1(t, \varepsilon) + C_2(t, \varepsilon) + C_3(t, \varepsilon) + C_4(t, \varepsilon), \text{ say,} \]
where \( C_{ij}(\varepsilon) \to 0 \text{ as } \varepsilon \downarrow 0. \) We know that \( \lim_{t \downarrow 0} C_1(t, \varepsilon) = 0 \) for fixed \( \varepsilon > 0 \) (Condition (vi)) and \( \lim_{t \downarrow 0} C_2(t, \varepsilon) = a^i(x_o) \frac{\partial f}{\partial x^i_o}, \) independently of small \( \varepsilon. \) By \((\ast)\) and Schwarz’s inequality \( \lim_{t \downarrow 0} C_4(t, \varepsilon) = 0, \) boundedly in \( t > 0. \) Also the left side has a finite limit as \( t \downarrow 0. \) So the difference
\[ \lim_{t \downarrow 0} C_3(t, \varepsilon) - \lim_{t \downarrow 0} C_3(t, \varepsilon) \]
can be made arbitrarily small by taking \( \varepsilon > 0 \) small. But by \((\ast)\), Schwarz’s inequality and (vi), the difference is independent of small \( \varepsilon > 0. \) Thus finite limit \( \lim_{t \downarrow 0} C_3(t, \varepsilon) \) exists independently of small \( \varepsilon > 0. \) Since we may choose \( F \in \mathcal{D}(A) \cap C^\infty \) such that
\[ \frac{\partial^2 F}{\partial x^i_o \partial x^j_o} \quad (i, j = 1, \ldots, n) \]
is arbitrarily near \( \alpha_{ij} \quad \alpha_{ij} \) being constants, it follows, by an argument similar to the one above that
finite limit
\[ \int_{d(x,x_o) \leq \varepsilon} (x^i - x_o^i)(x^j - x_o^j) P(t, x_o, dx) = b^{ij}(x_o) \]
eexists and
\[ \lim_{t \downarrow 0} C_3(t, \varepsilon) = b^{ij}(x_o) \frac{\partial^2 F}{\partial x^i_o \partial x^j_o}. \]
This completes the proof of the theorem. □

Remark. i) We have $b^{ij}(x) = b^{ij}(x)$ and

$$b^{ij}(x_o)\xi_i\xi_j \geq 0, (\xi_i \text{real}), \text{ for }$$

$$(x^j - x_o^j)(x^j - x_o^j)\xi_i\xi_j = \left(\sum (x^j - x_o^j)\xi_i\right)^2$$

ii) $b^{ij}(x)$ is a contravariant tensor:

$$\bar{b}^{ij} = b^{kl} \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial \bar{x}^j}{\partial x^l} (x^1, \ldots, x^n) \rightarrow (\bar{x}^1, \ldots, \bar{x}^n)$$

and

$$\bar{a}^m = a^s \frac{\partial \bar{x}^m}{\partial x^s} + b^{kl} \frac{\partial^2 \bar{x}^m}{\partial x^k \partial x^l}.$$ 

This follows from the equality

$$\bar{b}^{ij} \frac{\partial^2 f}{\partial \bar{x}_i \partial \bar{x}_o} + \bar{a}^m \frac{\partial f}{\partial \bar{x}^m} = b^{kl} \frac{\partial^2 f}{\partial x^k \partial x^l} + a^s \frac{\partial f}{\partial x^s}$$

[since each is $(Af)(x_o)$].
Part III

Regularity properties of solutions of linear elliptic differential equations
Lecture 13

The results proved in this part will be needed in the application of the semi-group theory to Cauchy’s problem.

1 Strong differentiability

Let $R$ be a subdomain of $E^m$. We denote by $C^\infty(R)$ the space of indefinitely differentiable functions in $R$ and by $\mathcal{D}^\infty(R)$ the space of $C^\infty$ functions in $R$ with compact support. We denote by $L_2(R)_{\text{loc}}$ the space of locally square summable functions in $R$, (i.e., functions in $R$ which are square summable on every compact subset of $R$). A function $u(x) \in L_2(R)_{\text{loc}}$ is said to be $k$-times strongly differentiable in $R$ (or of order $k$ in $R$) if for every subdomain $R_1$ of $R$ relatively compact in $R$ there exists a sequence $u_n(x)(=u_{n,R_1}(x))$ of $C^\infty$ functions in $R_1$, such that

$$\lim_{n \to \infty} \int_{R_1} |u - u_n|^2 dx = 0$$

and

$$\lim_{n,1 \to \infty} \int_{R_1} |D^{(s)}u_n - D^{(s)}u_1|^2 dx = 0 \quad \text{for } |s| \leq k.$$

Then there exists, for $|s| \leq k$, functions

$$u^{(s)}(x) = u^{(s)}_{R_1} \in L_2(R_1)$$

such that

$$\lim_{n \to \infty} \int_{R_1} |u^{(s)}(x) - D^{(s)}u_n(x)|^2 dx = 0.$$
\( u^{(s)}_{R_1}(x) \) is determined independently of the approximating sequence \( u_n \); for we have, for each \( C^\infty \) function \( \varphi \) with compact support in \( R_1 \)

\[
\int_{R_1} \varphi(x) u^{(s)}(x) dx = \lim_{n \to \infty} \int_{R_1} \varphi(x) D^{(s)} u_n(x) dx
\]

\[
= \lim_{n \to \infty} (-1)^{|s|} \int_{R_1} u_n(x) D^{(s)} \varphi(x) dx
\]

\[
= (-1)^{|s|} \int_{R_1} u(x) D^{(s)} \varphi(x) dx
\]

and \( C^\infty \) functions with compact support in \( R_1 \) are dense in \( L^2(R_1) \). It also follows that, for \( |s| \leq k \), there exists a function in \( L^2(R_1) \) denoted by \( \tilde{D}^{(s)} u(x) \), such that for each subdomain \( R_1 \) relatively compact in \( R \), \( \tilde{D}^{(s)} u(x) \) coincides with \( u^{(s)}_{R_1}(x) \) almost everywhere in \( R_1 \). \( \tilde{D}^{(s)} u(x) \) is called the strong derivative of \( u \) corresponding to the derivation \( D^{(s)} \).

### 2 Weak solutions of linear differential operators

Let

\[
L = \sum_{|\rho|=|\sigma|=\alpha} D^{(\rho)} a^{\rho\sigma} D^{(\sigma)}, \quad a^{\rho\sigma}(x) \in C^\infty(R), \quad a^{\rho\sigma} = a^{\sigma\rho} \quad \text{for} \quad |\sigma| = |\rho| = n,
\]

be a linear differential operator in \( R \) with \( C^\infty \) coefficients. Let \( f \in L^2(R)_{\text{loc}} \). A function \( u \in L^2(R)_{\text{loc}} \) will be said to be a weak solution of the equation \( Lu = f \) if for every \( \varphi \in \mathcal{D}^\infty(R) \) we have

\[
\int_R L^* \varphi u dx = \int_R \varphi f dx
\]

where \( L^* \) is the adjoint of \( L \):

\[
L^* = \sum_{|\rho|=|\sigma|=\alpha} (-1)^{|\rho|+|\sigma|} D^{(\sigma)} a^{\sigma\rho} D^{(\rho)}.
\]
3 Elliptic operators

Friedrichs - Lax - Nirenberg theorem: Let $L$ be elliptic in $R$ in the sense that there exists a constant $C_o > 0$ such that

$$\sum_{|\rho|=|\sigma|=n}\xi_1^{\rho_1}\cdots\xi_m^{\rho_m}d^{n-m}\omega_{m-n}(x)\xi_1^{\sigma_1}\cdots\xi_m^{\sigma_m} \geq C_o \left(\sum_{i=1}^m \xi_i^2\right)^n$$

for every $x \in R$ and every real vector $(\xi_1, \ldots, \xi_m)$. Then if $u_o$ is a weak solution of $Lu = f$ and if $f$ is of order $p$ in $R$, then $u_o$ is of order $2n + p$ in $R$.

Sobolev’s lemma: If $u_o(x)$ is of order $k$ in $R$, then, for $k > m/2 + \sigma$, $h_o(x)$ is equal almost everywhere (in $R$) to a function which is $\sigma$ times continuously differentiable.

Weyl-Schwartz theorem: Let $L$ be an elliptic operator in $R$, and $u_o$ a weak solution of $Lu = f$. If $f$ is indefinitely differentiable in $R$, then $u_o$ is almost everywhere equal to an indefinitely differentiable function in $R$.

This theorem is an immediate consequence of the Friedrichs Lax-Nirenberg theorem and Sobolev’s lemma.

4 Fourier Transforms:

For the proofs we need the following facts about Fourier transforms:

Plancherel’s theorem: Let $f(x) \in L_2(E^n)$, $x = (x_1, \ldots, x_n)$. Then the functions

$$\phi_n(y) = \int_{|x| \leq n} f(x) \exp(-2\pi i x \cdot y) \, dx \quad (x \cdot y = \sum x_i y_i)$$

converge in the $L_2$-norm to a function $\varphi(y_1, \ldots, y_n) \in L_2$ and the transformation $\mathcal{F}$ defined by $\mathcal{F}f = \varphi(y) = \lim_{n \to \infty} \int_{|x| \leq n} f(x) \exp(-2\pi i x \cdot y) \, dx$ is a unitary transformation of $L_2$ onto itself. (i.e., $(\mathcal{F}f, \mathcal{F}g) = (f, g)$, for $f, g \in L_2$ onto $L_2$). The inverse $\mathcal{F}^{-1}$ of $\mathcal{F}$ is given by

$$\mathcal{F}^{-1}\varphi(x) = \lim_{n \to \infty} \int_{|y| \leq n} \varphi(y) \exp(2\pi i y \cdot x) \, dy$$
\( \mathcal{F}(f) \) is called the Fourier transform of \( f \).

As regards the Fourier transform of the derivatives, we have: if \( f \in L^2(\mathbb{E}^m) \) is also in \( C^k(\mathbb{E}^m) \) and \( D^s f(x) \in L^2(\mathbb{E}^m) \) for \( |s| \leq k \), then

\[
(D^s f)(x) = \frac{\partial^{s_1 + \cdots + s_m}}{\partial x_1^{s_1} \cdots \partial x_m^{s_m}}, \quad |s| = \sum_{i=1}^m s_i,
\]

\( (\mathcal{F}D^s f)(y) = \prod_{j=1}^m (2\pi iy_j)^{s_j} \mathcal{F}(f)(y). \)

**Proof of Sobolev’s lemma:** Let \( R_1 \) be any relatively compact subdomain of \( R \) and \( \alpha(x) \) a \( C^\infty \) function with compact support in \( R \) such that \( \alpha(x) \equiv 1 \) on \( R_1 \). Since \( u_0 \) is assumed to be of order \( k \) there exists a sequence \( \{u_n\} \) of \( C^\infty \) functions in \( R_1 \) such that

\[
\lim_{n \to \infty} \sum_{|s| \leq k} \int_{R_1} |\hat{D}^s u_n - D^s u_n|^2 dx = 0.
\]

We have, using Leibniz’s formula,

\[
\lim_{n \to \infty} \sum_{|s| \leq k} \int_{R} |\hat{D}^s \alpha u_n - D^s \alpha u_n|^2 dx = 0.
\]

Let \( \tilde{u}_n \) (resp. \( \tilde{u}_o \)) denote the function in \( \mathbb{E}^m \) defined by:

\[
\tilde{u}_n(x) = \begin{cases} 
\alpha u_n(x), & x \in \text{Support of } \alpha \\
0, & x \in \mathbb{E}^m - \text{supp } \alpha;
\end{cases}
\]

similar definition for \( \tilde{u}_o(= \alpha u_o \text{ in supp. } \alpha) \). Since the Fourier transform is a unitary transformation, we have

\[
\lim_{n \to \infty} \| \mathcal{F}D^s \tilde{u}_n - \mathcal{F}D^s \tilde{u}_o \|_{L^2(\mathbb{E}^m)} = 0.
\]

But, as remarked earlier,

\[
(\mathcal{F}D^s \tilde{u}_n)(y) = (2\pi i)^s y_1^{s_1} \cdots y_m^{s_m} \hat{U}_n(y)
\]

where \( \hat{U}_n = \mathcal{F}\tilde{u}_n \); also since \( \mathcal{F} \) is unitary,

\[
\lim_{n \to \infty} \| \hat{U}_n - \hat{U}_o \|_{L^2(\mathbb{E}^m)} = 0, \text{ where } \hat{U}_o = \mathcal{F}(\tilde{u}_o).
\]
Therefore there exists a subsequence \( \{n'\} \) of \( \{n\} \) such that for almost all \( y \in E^m \)

\[
\lim_{n' \to \infty} \tilde{U}_{n'}(y) = \tilde{U}_o(y) \quad \text{(pointwise limit)}
\]

\[
\lim_{n' \to \infty} \tilde{U}_{n'}(y)x_1^{s_1} \cdots x_m^{s_m}(2\pi i)^{|l|} = \tilde{U}_o[y] = \tilde{U}_o^{(s)}(y)
\]

where

\[
\tilde{U}_o^{(s)} = \mathcal{F}\tilde{D}^{(s)}\tilde{u}_o.
\]

Thus for almost all \( y \in E^m, \tilde{U}_o(y)x_1^{s_1} \cdots x_m^{s_m}(2\pi i)^{|l|} = \tilde{U}_o^{(s)}(y), |s| \leq k. \)

We shall now show that \( \tilde{U}_o(y) \cdot \tilde{y}_1^q \cdots \tilde{y}_m^q \) is integrable on \( E^m \) provided \( k > \frac{m}{2} + \sigma \), where \( \sigma = \frac{m}{2} \sum q_j \). We have

\[
\tilde{U}_o(y)\tilde{y}_1^q \cdots \tilde{y}_m^q = \frac{\tilde{y}_1^q \cdots \tilde{y}_m^q}{1 + |\sum_{i=1}^m \tilde{y}_i^{2k/2}\tilde{U}_o(y)} \left(1 + \sum_{i=1}^m \tilde{y}_i^{2k/2}\right).
\]

Now, in polar coordinates

\[
dy = dy_1 \cdots dy_m = r^{m-1} dr \Omega_{m-1}
\]

(\( \Omega_{m-1} \) is the surface of unit sphere in \( E^m \)). So \( \frac{\tilde{y}_1^q \cdots \tilde{y}_m^q}{1 + |\sum_{i=1}^m \tilde{y}_i^{2k/2}\tilde{U}_o(y)} \) is square integrable in \( |z| > a(Z \in E^m) \) if \( 2|q| - 2k + m - 1 < -1 \), i.e., if \( k > \frac{m}{2} + \sigma \). Already we know that \( U_o(y)(1 + \sum_{i=1}^m \tilde{y}_i^{2k/2} \) is square integrable in \( |z| > a \). So \( U_o(y)\tilde{y}_1^q \cdots \tilde{y}_m^q \), begin the product of two square integrable functions, is integrable in \( |z| > a \). We see also that \( U_o(y)\tilde{y}_1^q \cdots \tilde{y}_m^q \) is integrable in \( |z| \leq a \).

Thus if \( k > \frac{m}{2} + |q|, U_o(y)\tilde{y}_1^q \cdots \tilde{y}_m^q \) is integrable over \( E^m \).

Suppose \( k > \frac{m}{2} + \sigma, (\sigma > 0 \text{ integer}). \) Then \( \tilde{U}_o(y) \in L_2 \cap L_1 \) so that \( (\mathcal{F}^{-1} \tilde{U}_o)(y) = \int E^m \tilde{U}_o(y) \exp(2\pi iy_x)dy, a.e \) on \( E^m; \) i.e., \( \tilde{u}_o(x) = \int E^m \tilde{U}_o(y) \exp(2\pi iy_x)dy, a.e \) on \( E^m. \)
Let $|q| \leq \sigma (k > \frac{m}{2} + \sigma)$; then

\[
D^{(q)}_x \left\{ \hat{U}_o(y) \exp(2\pi i y.x) \right\} = \hat{U}_o(y) \prod_{j=1}^{m} (2\pi iy_j)^{q_j} \exp 2\pi iy.x
\]

and

\[
\left| \hat{U}_o(y) \prod_{j=1}^{m} (2\pi iy_j)^{q_j} \right| \exp 2\pi iy.x \leq \left| \hat{U}_o(y) \prod_{j=1}^{m} (2\pi iy_j)^{q_j} \right|
\]

and $\left| \hat{U}_o(y) \prod_{j=1}^{m} (2\pi iy_j)^{q_j} \right|$ is a function independent of $x$ and summable (as a function of $y$) over $E^m$. Therefore $D^{(q)}(x)\hat{u}_o(x)$ exists and $D^{(q)}\hat{u}_o(x) = \int_E \hat{U}_o(y) \prod_{j=1}^{m} (2\pi iy_j)^{q_j}(\exp 2\pi iy.x)dy$.

This representation also shows that $D^{(q)}\hat{u}_o(x)$ is continuous. Thus $\hat{u}_o(x)$ is $\sigma$-times continuously differentiable; so $u_o(x)$ is $\sigma$-times continuously differentiable in $R_1$. 
Lecture 14

1 Garding’s inequality

For the proof of the Friedrichs-Lax-Nirenberg theorem, we need Garding’s inequality. Let $R_1$ be a relatively compact subdomain of $R$ and let $L$ be a linear elliptic differential operator in $R$. There exist $\alpha > 0$ and $\delta > 0$ such that for $\varphi \in \mathcal{D}^\infty(R_1)$,

$$\langle \varphi + \alpha(-1)^n L^* \varphi, \varphi \rangle \geq \delta \|\varphi\|_n^2$$

where

$$\|\varphi\|_n = \int_{R_1} \sum_{|\beta| \leq n} |D^\beta \varphi|^2 \, dx.$$

Before proving the theorem, we prove a preliminary proposition.

**Proposition.** (i) Define for $\varphi \in \mathcal{D}^\infty(R_1)$

$$\|\varphi\|_j^2 = \sum_{|\beta| = j} \int_{R_1} |D^\beta \varphi|^2 \, dx.$$

Then for $j < n$ there exists a positive constant $e^{jn}$ such that

$$\|\varphi\|_j \leq e^{jn} \|\varphi\|_n$$

(ii) \( \lim_{\alpha \to 0} \sup_{\varphi \in \mathcal{D}^\infty(R_1)} \left\{ \frac{\alpha \|\varphi\|_{n+1}^2}{\|\varphi\|_n^2 + \alpha \|\varphi\|_n^2} \right\} \)

93
(iii) There exist positive constants $\mu$ and $\mu'$ such that for $\varphi \in D^\infty(R_1)$

$$\sum_{\|x\|=\|\varphi\|=\|\psi\|} (D^{\sigma}d^{\rho_\sigma}D^{\psi},\varphi) \geq \mu\|\varphi\|^2_2 - \mu'\|\varphi\|_1\|\varphi\|_\infty$$

80 Proof. (i) Let

$$\tilde{\varphi}(x) = \begin{cases} \varphi(x) & x \in R_1 \\ 0 & x \in E^m - R_1 \end{cases}$$

Then

$$\tilde{\varphi}(x) = \tilde{\varphi}(x_1, \ldots, x_m) = \int_{-\infty}^{\infty} \partial(x_1, \ldots, x_{s-1}, l, x_{s+1}, \ldots, x_m) \frac{\partial}{\partial t} dt$$

Hence by Schwarz’s inequality

$$|\tilde{\varphi}(x)|^2 \leq L \int_{-\infty}^{\infty} \left\| \frac{\partial \tilde{\varphi}}{\partial x_s} \right\|^2 dx_s,$$

where $L$ is the diameter of $R_1$. So

$$\int_{R_1} |\varphi|^2 dx = \int_{R_1} \tilde{\varphi}^2 dx$$

$$\leq L \int_{R_1} dx_1 \cdots dx_m \left\{ \int_{-\infty}^{\infty} \left| \frac{\partial \tilde{\varphi}}{\partial x_s} \right|^2 dx_s \right\}$$

$$= L^2 \int_{R_1} \left| \frac{\partial \varphi}{\partial x_s} \right|^2 dx$$

Therefore

$$\|\varphi\|^2_2 \leq L^2 \left\| \frac{\partial \varphi}{\partial x_s} \right\|^2_2.$$  

By repeated application of this inequality we get (i).

(ii) Since

$$\mathcal{F}D^{(s)}\tilde{\varphi}(y) = \prod_{j=1}^{m} (2\pi i y_j)^{s_j} \phi(y), (\phi = \mathcal{F}\tilde{\phi})$$
1. Garding’s inequality

and \( \mathcal{F} \) is a unitary transformation in \( L_2 \), we obtain

\[
\|\|\phi\|\|^2_l = \sum_{|s|=1}^{E_m} \int |\mathcal{F}D^{(s)}\phi|^2 dx
\]

\[
= (2\pi)^{2l} \sum_{|s|=1}^{E_m} \int \prod_{j=1}^{m} y_j^{2s_j} |\Phi(y)|^2 dy.
\]

Since

\[
\frac{\alpha \sum_{|s|=n-1} \prod_{j=1}^{m} y_j^{2s_j}}{1 + \alpha \sum_{|s|=n} \prod_{j=1}^{m} y_j^{2s_j}}
\]

tends to zero uniformly in \( y \) as \( \alpha \downarrow 0 \).

(iii) is proved.

(iv) When \( a^{n_1,\ldots,n_m}(x) \) with

\[
\sum n_i = \sum n'_i = n
\]

are constant we have by partial integration and Fourier transform

\[
\sum_{|\rho|=|\sigma|=n} D^{(\rho)} a^{\rho,\sigma} D^{(\sigma)} \phi, \phi = \sum_{|\rho|=|\sigma|=n} (-1)^n a^{\rho,\sigma} (D^\rho \phi, D^\sigma \phi)
\]

\[
= \int \sum_{|\rho|=|\sigma|=n} (2\pi)^{2n} y_1^{2\rho_1} \cdots y_m^{2\rho_m} a^{\rho_1,\ldots,\rho_m,\sigma_1,\ldots,\sigma_m} y_1^{2\sigma_1} \cdots y_m^{2\sigma_m} \geq \text{Const} \int \sum_{|\rho|=n} |y_1^{2\rho_1} \cdots y_m^{2\rho_m}|^2 |D^{(\rho)} \phi(y)|^2 dy |D^{(\sigma)} \phi|^2 dy
\]

(making use of the ellipticity)

\[
= \text{Const} \int \sum_{|\rho|=n} |D^{(\rho)} \phi|^2 dx
\]

\[
\geq \text{Const} \|\|\phi\|\|^2_l.
\]

If \( a^{\rho,\sigma}(x), (|\rho| = |\sigma| = n) \) are non-constant, put

\[
\varepsilon = \sup_{\rho,\sigma;x',x'' \in \mathbb{R}_1} |a^{\rho,\sigma}(x') - a^{\rho,\sigma}(x'')|.
\]
Note that $\varepsilon$ may be taken to be arbitrarily small if we choose $R_1$ sufficiently small. Let $x^o$ be a fixed point of $R_1$. Put $a^{\rho,\sigma}(x^o) = a_0^{\rho,\sigma}$.

Let $\varphi \in \mathcal{D}^{\infty}(R_1)$. We have

$$\sum_{|\rho|=|\sigma|=n} (-1)^{n}(d^{\rho,\sigma} D^\rho \varphi, D^\sigma \varphi)$$

$$= \sum_{|\rho|=|\sigma|=n} (-1)^{n}(d_0^{\rho,\sigma} D^\rho \varphi, D^\sigma \varphi)_0 + \sum_{|\rho|=|\sigma|=n} (-1)^{n}[(d^{\rho,\sigma} - d_0^{\rho,\sigma}) D^\rho \varphi, D^\sigma \varphi)_0];$$

$$| \sum_{|\rho|=|\sigma|=n} (-1)^{n}(d^{\rho,\sigma} - d_0^{\rho,\sigma}) D^\rho \varphi, D^\sigma \varphi)_0 | \leq \varepsilon \sum_{|\rho|=|\sigma|=n} ||D^\rho \varphi||_0 ||D^\sigma \varphi||_0$$

$$\leq \text{Const} ||\varepsilon||^2.$$  

So

$$\sum_{|\rho|=|\sigma|=n} (-1)^{n}(d^{\rho,\sigma} D^\rho \varphi, D^\sigma \varphi) \geq C_1 ||\varphi||^2 - \text{Const} ||\varepsilon||^2$$

$$\geq C_3 ||\varphi||^2 (C_3 > 0).$$

if we choose, $R_1$ sufficiently small. This result enables us to deduce (iii) for the general case. For any $\eta > 0$, $R_1$ can be covered by a finite number, say $N$, of open spheres $S_1, S_2, \ldots, S_N$ of radius $\eta/2$. Let $S'_i$ be the sphere of radius $\eta$ concentric with $S_i$. Let $\varphi_i(x) \in C^{\infty}(E^m)$ satisfy

$\varphi_i(x) > 0$ for $x \in S_i$, $\varphi_i(x) = 0$ for $x \notin S'_i$ and $\varphi_i(x) \geq 0$ for $x \in E^m$.

Then

$$h_i(x) = (\varphi_i(x)/\sum_{j=1}^{N} \varphi_j(x))^{1/2}$$

satisfies $h_i(x) \in C^{\infty}(R_1) \setminus 0$ and $\sum_{j=1}^{N} h_j(x) \equiv 1$ or $1 + R$.

Thus

$$(-1)^{n} \sum_{|\rho|=|\sigma|=n} (d^{\rho,\sigma} D^\rho \varphi, D^\sigma \varphi)_0$$
1. Garding’s inequality

\[ \sum_{j=1}^{N} A_j = \sum_{j=1}^{N} (-1)^n \sum_{|\rho|=|\sigma|=n} (d^{\rho,\sigma} h_j D^\rho \varphi, h_j D^{(\sigma)} \varphi) \]

is such that

\[ A_j = (-1)^n \left\{ \sum_{|\rho|=|\sigma|=n} (d^{\rho,\sigma} D^\rho h_j \varphi, D^{(\sigma)} h_j \varphi) - R_j \right\} \]

where, by Leibnitz’s formula,

\[ R_j = \sum_{|\rho'|<n} (d^{\rho',\sigma'} D^{\rho'} \varphi, D^{(\sigma')} \varphi) \]

with bounded functions \( C^{\rho',\sigma'} \). Thus, by Schwarz’s inequality,

\[ |R_j| \leq a_j ||\varphi||_{l_n-1} ||\varphi||_n \quad (a_j = \text{constant} > 0). \]

For sufficiently small \( \eta > 0 \), we have, by the result obtained already,

\[ (-1)^n \sum_{|\rho|=|\sigma|=n} (d^{\rho,\sigma} D^\rho \varphi, D^{(\sigma)} \varphi) \]

\[ \leq \sum_{j=1}^{n} (\lambda_j ||h_j||^2_n - a_j ||\varphi||_{l_n-1} ||\varphi||_n) \quad (\lambda_j = \text{Const} > 0) \]

Moreover, we have, by the same reasoning as above,

\[ ||h_j \varphi||^2_n \geq \int_R h_j^2(x) \sum_{|\rho|=n} |D^{(\rho)} \varphi(x)|^2 dx - b_j ||\varphi||_{l_n-1} ||\varphi||_n \]

with constant \( b_j > 0 \). Therefore, by putting

\[ \lambda = \min(\lambda_j), \sum_{j=1}^{N} (\lambda_j b_j + a_j) = \lambda' \]

we have

\[ (-1)^n \sum_{|\rho|=|\sigma|=n} (d^{\rho,\sigma} D^{(\rho)} \varphi, D^{(\sigma)} \varphi) \geq \lambda ||\varphi||^2_n - \lambda' ||\varphi||_{l_n-1} ||\varphi||_n \]
Proof of Garding’s inequality: We have, by integration by parts and from part (iii) of the above proposition, for $\alpha > 0$.

$$
(\varphi + \alpha(-1)^n L^* \varphi, \varphi)_o \geq (\varphi, \varphi)_o + \alpha(\mu \| \varphi \|_2^2 - \mu' \| \varphi \|_{n-1} \| \varphi \|_n) + \sum_{|\sigma|<n \sigma \leq n} (C^0 \sigma D^{(\sigma)} \varphi, D^{\sigma} \varphi)_o
$$

where $C^0 \sigma$ are bounded $C^\infty$ functions in $R_1$. Then by (i) and Schwarz’s inequality

$$
(\varphi + \alpha(-1)^n L^* \varphi, \varphi) \geq \| \varphi \|_o^2 + \alpha\{\mu\| \varphi \|_n^2 - \eta \| \varphi \|_{n-1} \| \varphi \|_n\}
$$

with some positive constant $\eta$. Hence for any $\tau > 0$ we have, remembering

$$
\| \varphi \|_2^2 = \| \varphi \|_n^2 - \sum_{s<n} \| \varphi \|_s^2
$$

and using (i),

$$
(\varphi + \alpha(-1)^n L^* \varphi, \varphi) \geq \| \varphi \|_0^2 + \alpha \left\{ \mu \| \varphi \|_n^2 - \mu'' \| \varphi \|_{n-1}^2 - \frac{\eta}{2} \left( \| \varphi \|_{n-1}^2 + \| \varphi \|_n^2 + \| \varphi \|_n^2 \right) \right\}
$$

Then by taking $\tau^{-1} > 0$ so small that $(\mu - \eta/2\tau^{-1}) > 0$ and $\alpha > 0$ sufficiently small we obtain Garding’s inequality by (ii).
Lecture 15

1 Proof of the Friedrichs - Lax - Nirenberg theorem

To prove the Friedrichs - Lax - Nirenberg theorem, we need three lemmas:

**Lemma 1.** If $u_0$ is of order $i$ in $R_1$ and if $\hat{D}(s)u_0$ is of order $j$ in $R$, for all $s$ with $|s| \leq i$, then $u_0$ is of order $i + j$ in $R$. If $u_0$ is of order $i + j$ in $R$, then $\hat{D}(s)u_0$ is of order $j$ for $|s| \leq i$.

**Lemma 2.** Let $R_1$ be a relatively compact subdomain of $R$ and let $u_0 \in L^2(R_1)$. Then for any positive integer $s$

\[
(I + (-\Delta)^s)h = u_0 \quad (\Delta \text{ is the Laplacian})
\]

has weak solution of order $2s$ in $R_1$.

**Lemma 3.** Let $u_0 \in L^2(R_1)$ be of order $n$ in $R_1$ and

\[
|\langle L^s \varphi, u_0 \rangle| \leq \text{Const} \|\varphi\|_{n-1}, \quad \text{for all} \ \varphi \in \mathcal{D}(R_1)
\]

Then $u_0$ is of order $n + 1$ in $R_1$.

Assuming these lemmas for a moment, we shall give a *Proof of the Friedrichs - Lax - Nirenberg theorem*. 84
First Step 1. If \( u_o \in L_2(R_1) \) is of order \( n \) in \( R_1 \) and satisfies \( \|L^* \varphi, u_o \| \leq \text{Const} \| \varphi \|_{n-j} \) for all \( \varphi \in \mathcal{D}^{\infty}(R_1) \), then \( u_o \) is of order \( n + j \) in \( R_1 \). This is proved by induction on \( j \). The result is true for \( j = 1 \) (Lemma 3). Let us assume that \( j > 1 \) and that the result is true for \( j - 1 \). Suppose

\[
\|L^* \varphi, u_o \| \leq \text{Const} \| \varphi \|_{n-j};
\]

since \( \| \varphi \|_{n-j} \leq \| \varphi \|_{n-(j-1)}, u_o \) is of order \( (n + j - 1) \) in \( R_1 \) by the inductive assumption. For any first order derivation \( D \), we have \( \|L^* D \varphi, u_o \| \text{Const} \| D \varphi \|_{n-j} \leq \text{Const} \| \varphi \|_{n-j+1} \). Since \( u_o \) is of order \( n + 1 \), we have

\[
(L^* D \varphi, u_o) = \sum_{|\rho| \leq n} \left( (-1)^{|\rho|} D^{\rho} \varphi, a^{\rho} D^{\varphi} D \phi, u_o \right)
\]

\[
= \sum \left( (-1)^{|\rho|} D^{\rho} \varphi, a^{\rho} D^{\phi} D \phi, u_o \right)
\]

\[
= \sum \left( (-1)^{|\rho|+1} D^{\rho} \varphi, D(a^{\rho} D^{\phi}, D \phi, u_o) \right)
\]

\[
= \sum \left( (-1)^{|\rho|+1} D^{\rho} \varphi, D(a^{\rho} D^{\phi}, D \phi, u_o) \right)
\]

\[
= \sum \left( (-1)^{|\rho|+1} D^{\rho} \varphi, D(a^{\rho} D^{\phi}, D \phi, u_o) \right) - (L^* \varphi, D \phi, u_o).
\]

Since \( u_o \) is of order \( (n + j - 1) \) we see by partial integration that

\[
\|L^* \varphi, D \phi, u_o \| \leq \|L^* D \varphi, u_o \| + \text{Const} \| \varphi \|_{2n} - (n + j - 1)
\]

\[
\leq \text{Const} \| \varphi \|_{n-(j-1)}
\]

By Lemma 1 \( D \phi, u_o \) is of order \( \geq n + j - 1 \geq n + j - 2 \geq n \) (as \( j \geq 2 \)). Hence by the induction assumption \( D \phi, u_o \) is of order \( n + j - 1 \). So, by lemma 1 \( u_o \) is of order \( n + j \).

Second Step 1 (Friedrich’s theorem). Let \( u_o \in L_2(R_1) \) be a weak solution of \( Lu = f \) and \( f \) be order \( p \) in \( R_1 \). If \( u_o \) is of order \( n \) in \( R_1 \), then \( u_o \) is of order \( 2n + p \) in \( R_1 \).

Proof. This holds for \( p = 0 \). For, from \( (L^* \varphi, u_o)_o = (\varphi, f)_o \), we have

\[
\|L^* \varphi, u_o \|_o \leq \text{Const} \| \varphi \|_o = \text{Const} \| \varphi \|_{n-n}.
\]
1. Proof of the Friedrichs - Lax - Nirenberg theorem

So, by the first step $u_o$ is of order $n + n = 2n$. Suppose $p = 1$. We have, as above,

\[ (L^* \varphi, \tilde{D} u_o) = -(D L^* \varphi, u_o) = (-1)^{\lceil \frac{n}{2} \rceil + 1} (D^\sigma D^{\alpha, \sigma} D^\varphi, u_o) \]

\[ = (-1)^{\lceil \frac{n}{2} \rceil + 1} (D^\sigma D^{\alpha, \sigma} D^\varphi, u_o) \]

where $L'$ is a differential operator of degree $2n$.

\[ (L^* \varphi, \tilde{D} u_o) = (D \varphi, f)_o + (\varphi, \tilde{L}' u_o) \]

\[ = -(\varphi, D f) + (\varphi, \tilde{L}' u_o) \]

(since $f$ is of order 1 at least; the case $p = 0$ is already proved). Thus

\[ |(L^* \varphi, D u_o)_o| \leq \text{Const} \| \varphi \|_o = \text{Const} \| \varphi \|_{n-n} \]

and $\tilde{D} u_o$ is of order $2n - 1 \geq n$. So by the first step, $\tilde{D} u_o$ is of order $n + n = 2n$. By Lemma 1, $u_o$ is of order $2n + 1$. For $p > 1$, we may repeat the argument. \( \square \)

**Third Step I.** Let $u_o \in L_2(R_1)$ be a weak solution of $L u = f$ and $f$ be of order $p$ in $R_1$. Then $u_o$ is of order $2n + p$ in $R_1$.

**Proof.** Let $h_o$ of order $2n$ be a weak solution of

\[ (I + (-A)^n)h = u_o. \]

$h_o$ exists by Lemma 2. Then $h_o$ of order $2n$ is a weak solution of

\[ L(I + (-\triangle)^n)h = f; \]

$L(I + (-\triangle)^n)$ is an elliptic operator of order $4n$. $f$ being of order $p$, $h_o$ is of order $4n + p$, by the second step. Hence, by Lemma 1

\[ u_o = (I + (-\triangle)^n)h_o \]

is of order $4n + p - 2n = 2n + p$. \( \square \)
Lecture

1 Proof of Lemma 3

Let $R$ be a bounded domain of $E^m$. Let $u_0$ of order $n$ satisfy

$$|(L^*\varphi, u_0)_{L^2(\cdot)}| \leq \text{Const} \left|\varphi\right|_{n-1}$$

for all $\varphi \in D^\infty(R)$

Let $R_2 \subset R_1 \subset R$, $R_2, R_1$ being subdomains, such that the closure of $R_1$ in $R$ is compact. Let $\zeta \in D^\infty$ with $\zeta(x) = 1$ on $R_2$. Let

$$v^h(x) = \frac{v(x^h) - v(x)}{h}, x^h = (x_1 + h, x_2, \ldots, x_m),$$

$h$ sufficiently small. Then, as will be proved below,

$$\|v^h\|_{L^2} \leq \text{Const} \quad (\text{for all sufficiently small } h).$$

Since the Hilbert space $H_n(R)$ (completion of $D^\infty(R)$ by the norm $\|\cdot\|_n$) is locally weakly compact, there exists a sequence $\{h_i\}$ with $\lim_{i \to \infty} h_i = 0$ such that for $|k| \leq n$

$$\text{weak lim}_{i \to \infty} v^{h_i} = \hat{v}$$
$$\text{weak lim}_{i \to \infty} D^{k}v^{h_i} = v^{(k)}$$

exist in $L_2(R_1)$. We shall show that

$$\hat{v} = \hat{D}_1 v (D_1 = \partial/\partial x_1)$$

103
\[ v^{(k)} = \tilde{D}_1 \tilde{D}_v^{(k)} = \tilde{D}^{(k)} \tilde{D}_1 v \]

proving that \( \tilde{D}_1 v \) is of order \( n \) in \( R_1 \). Similarly \( \tilde{D}_i v (i = 2, \ldots, m) \) will be of order \( n \) in \( R_1 \) and hence \( u \) is of order \( n + 1 \) in \( R \). That \( \hat{v} = \tilde{D}_1 v \) may be proved as follows: For any \( \varphi \in \mathcal{D}^\infty (R_1) \) we have, \( \theta \) being a real number such that \( 0 < \theta < 1 \),

\[
\begin{align*}
(\varphi, \hat{v})_o &= \lim_{i \to \infty} (\varphi, v^h)_o \\
&= \lim_{i \to \infty} (\varphi^{-h}, v) \\
&= \lim_{i \to \infty} (\varphi(x^{-\theta h}), v(x))_o \\
&= \lim_{i \to \infty} (\varphi(x(-\theta h^i)), \tilde{D}_1 v(x))_o \\
&= (\varphi, D_1 v)_o.
\end{align*}
\]

We have also

\[ (D^k v)^h = D^k v^h \]

and thus, in \( L_2 \),

\[
v^k = \text{weak lim}_{i \to \infty} \tilde{D}^k v^h = w - \text{lim}_{i \to \infty} (\tilde{D}^k v)^h = \tilde{D}_1 D^{(k)} v.
\]

We prove that

\[ \|v^h\|_m \leq \text{Const} \text{ (for all small } h). \]

We shall make use of Garding’s inequality for the \( 2n \) order elliptic differential operator \( L^* \): there exist constants \( C_1, C_2 \) and \( C_3 \) such that

\[
C_1 \|\varphi\|^2_\omega \leq (L^* \varphi, \varphi)_o + C_2 \|\varphi\|^2_o \\
(L^* \varphi, \psi)_o \leq C_3 \|\varphi\|_m \|\psi\|_o, \quad \varphi, \psi \in \mathcal{D}^\infty (R).
\]

Now,

\[
(L^* \varphi, v^h)_o = (-1)^{\ell} (D^\ell \varphi, a^{\ell \sigma} \tilde{D}^{(\sigma)} (\xi u_0)^h)_o \\
= (-1)^{\ell} (D^\ell \varphi, a^{\ell \sigma} (\tilde{D}^{(\sigma)} \xi u_0)^h)_o
\]

where

\[ u = \tilde{D}_1 v \]
1. Proof of Lemma 3

\[ = (-1)^{\rho_1} (D^{(\rho)} \varphi, a^{\rho,\sigma} (\zeta \cdot \bar{D}^{(\sigma)} u_o)_o) \]

\[ + (-1)^{\rho_1} C^{\sigma,\rho,\sigma} (D^\rho \varphi, a^{\rho,\sigma} [D^{\rho'} \zeta \cdot \bar{D}^{(\sigma')} u_o]_o^h) \quad (|\sigma'| \geq 1) \]

by applying the Leibnitz formula.

On the other hand, we have, for any function \( w \) of order \( j \) in \( R \) with support completely interior to \( R \)

\[ \| w^h \|_{j-1, R_1} \leq \| w \|_j \quad \text{for sufficiently small } |h|, \text{ because, for any approximating functions } \{u_i\} \leq C^{\infty}(R) \]

\[ \| w^h \|_{j-1, R_1} = \lim_{i \to \infty} \| u_i^h \|_{j-1, R_1} \]

\[ = \lim_{i \to \infty} \| u_{i_j} (x^{(ih)}) \|_{j-1, R_1} \]

\[ \leq \| w \|_{j,R} = \| w \|_j. \]

Thus the absolute value of the second term on the right of (*) is by Schwarz’s inequality \( \leq \text{Const } \| \varphi \|_m \| u \|_n = \text{Const } \| \varphi \|_m \). Since \( (ef)^h(x) = e^h(x) f(x^h) - e(x) f^h(x), \)

we have

\[ (-1)^{\rho_1} (D^\rho \varphi, a^{\rho,\sigma} (\zeta \cdot \bar{D}^{(\sigma)} u_o)_o^h) \]

\[ = (-1)^{\rho_1} (D^\rho \varphi, [(a^{\rho,\sigma} \zeta \cdot \bar{D}^{(\sigma)} u_o)_o^h - (a^{\rho,\sigma} \zeta \cdot \bar{D}^{(\sigma)} u_o(x^h))]_o) \]

\[ = (-1)^{\rho_1} (D^\rho \varphi, a^{\rho,\sigma} \zeta \cdot \bar{D}^{(\sigma)} u_o)_o \]

\[ + (-1)^{\rho_1+1} (D^\rho \varphi, (a^{\rho,\sigma} \zeta \cdot \bar{D}^{(\sigma)} u_o(x^h))_o \]

The absolute value of the second term on the right is

\[ \leq \text{Const } \| \varphi \|_m. \]

We have also

\[ (-1)^{\rho_1} ((D^\rho \varphi)^{-h}, a^{\rho,\sigma} \zeta \cdot \bar{D}^{(\sigma)} u_o)_o \]

\[ = (-1)^{\rho_1} (D^\rho \varphi^{-h}, a^{\rho,\sigma} \zeta \cdot \bar{D}^{(\sigma)} u_o)_o \]

\[ = (-1)^{\rho_1} (a^{\rho,\sigma} \zeta \cdot D^\rho \varphi^{-h}, \bar{D}^{(\sigma)} u_o)_o. \]
\[ = (-1)^{\text{deg}} (a^{\alpha \sigma} D^\rho \zeta \varphi^{-h}, \tilde{D}^\sigma u_o) \]
\[ = (-1)^{\text{deg}} C^{\varphi^*} (a^{\alpha \sigma} (D^\rho \zeta D^{(\rho') \varphi^*})^{-h}, \tilde{D}^\sigma u_o) (|\varphi'| \geq 1). \]

The absolute value of the second term on the right is
\[ \leq \text{Const} \| \varphi^{-h} \|_{n-1} \leq \text{Const} \| \varphi \|_{n-1}. \]

Therefore, by applying the original hypothesis,
\[ |(L^* \varphi, (\zeta u_o)^h) | \leq |(L^* \zeta \varphi^{-h}, u_o) | + \text{Const} \| \varphi \|_n \]
\[ \leq \text{Const} \| \zeta \varphi^{-h} \|_{n-1} + \text{Const} \| \varphi \|_n \]
\[ \leq K \| \varphi \|_n, \text{ } K \text{ a positive constant}. \]

Thus letting \( \varphi \) tend, in \( \| \| \), to \( (\zeta u_o)^h \), we have
\[ C_1 \| (\zeta u_o)^h \|_n^2 \leq K \| (\zeta u_o)^h \|_n + C_2 \| (\zeta u_o)^h \|_n \]

Since \( \| (\zeta u_o)^h \|_n \leq \text{Const} \| \zeta u_o \|_n \), the right hand side being independent of \( h \), we must have
\[ \| (\zeta u_o)^h \|_n \leq \text{Const} ( \text{ independent of } h). \]
Lecture 17

1 Proof of Lemma 2

We define

\[ \tilde{u}_o(x) = \begin{cases} u_o(x) & \text{if } x \in R_1 \\ 0 & \text{if } x \in E^m - R_1. \end{cases} \]

Let \( U_o(y) = (\mathcal{F} \tilde{u}_o)(y) \). Then

\[ h_o(x) = \mathcal{F}^{-1} \frac{U_o(Y)}{1 + (\sum_{j=1}^{m} (2\pi y_j)^2)^x} (x) \]

satisfies the conditions of the lemma. In the first place,

\[ h_o(x) = \int_{|y| \leq n} \frac{U_o(y)}{1 + (\sum_{j=1}^{m} (2\pi y_j)^2)^x} \exp \left(2\pi \sqrt{-1} y x \right) dy \]

is \( C^m(E^m) \). For, since \( U_o(y) \in L_2(E^m) \),

\[ \frac{U_o(y)}{1 + (\sum_{j=1}^{m} (2\pi y_j)^2)^x} \prod_{j=1}^{m} (2\pi \sqrt{-1} y_j)^{k_j} \exp \left(2\pi \sqrt{-1} y x \right) \]

is, for any set of integers \( k_j \geq 0 \), integrable over \(|y| \leq n\) and majorised uniformly in \( x \) by a summable function (in \( y \)). So

\[ \frac{\partial^{k_1 + \ldots + k_m}}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} h_o(x) = \int_{|y| \leq n} \left\{ \frac{U_o(y) \prod_{j=1}^{m} (2\pi \sqrt{-1} y_j)^{k_j}}{1 + (\sum_{j=1}^{m} (2\pi y_j)^2)^x} \right\} \exp(2\pi \sqrt{-1} y x) dy \]
Moreover, for $|k| \leq 2s$, the function under the curly brackets \{\cdots\} is in $L_2(E^m)$, so that for $|k| \leq 2s, D^{(k)}h_n(E^x)$ converges in $L_2(E^m)$. Therefore $h_n(x)$ is of order $2s$ in $E^m$.

Next for any $\varphi \in D^\infty(E^m)$, we have, by partial integration,

$$\int_{E^m} (I + (-\Delta)^s) \varphi(x) h_n(x) \, dx = \lim_{n \to \infty} \int_{E^m} (I + (-\Delta)^s) \varphi(x) h_n(x) \, dx$$

$$= \lim_{n \to \infty} \int_{E^m} \varphi(x) (I + (-\Delta)^s) h_n(x) \, dx$$

$$= \int_{E^m} \varphi(x) (\mathcal{F}^{-1} U_o) (x) \, dx$$

This proves that $h_n$ is a weak solution in $E^m$ of $(I + (-\Delta)^s) h = \tilde{u}_o = \mathcal{F}^{-1} U_o$. Thus $h_n$ is a weak solution in $R_1$ of $(I + (-\Delta)^s) h = u_o$.

## 2 Proof of Lemma 1

In the proof of Lemma 1 we have to make use of the notion of “regularisation” or “mollifiers”. Let $j(x) \in C^\infty(E^m)$ such that

i) $j(x) \geq 0$,

ii) $j(x) = 0$, for $|x| \geq 1$

iii) $\int_{E^m} j(x) \, dx = 1.$

Let for $\varepsilon > 0$

$$j_\varepsilon(x) = \varepsilon^{-n} j(x/\varepsilon)$$

We have then

i) $j_\varepsilon(x) \geq 0$.

ii) $j_\varepsilon(x) = 0$, for $|x| \geq \varepsilon$

iii) $\int_{E^m} j_\varepsilon(x) \, dx = 1.$
Let \( R_1 \) be a relatively compact subdomain of \( R \subset E^m \) and \( u(x) \in L^2(R_1) \). Let \( R_2 \) be a subdomain relatively compact in \( R_1 \). Let \( d > 0 \) be the distance between \( R_2 \) and the boundary of \( R_1 \). Let \( \varepsilon > 0 \) be such that \( \varepsilon < d \). For \( x \in R_2 \), define

\[
(J_\varepsilon u)(x) = \int_{R_1} j_\varepsilon(x - y)u'(y) \, dy.
\]

\((J_\varepsilon u)(x)\) is called regularisation of \( u(x) \) and the operators \( J_\varepsilon \) are called mollifiers. Let

\[
\|v\|_{o,R_i}^2 = \int_{R_i} |v|^2 \, dx.
\]

We then have

i) \( \| J_\varepsilon u \|_{o,R_2} \leq \| u \|_{o,R_1} \)

ii) \( \lim_{\varepsilon \downarrow 0} \| J_\varepsilon u - u \|_{o,R_2} = 0 \)

iii) \( (J_\varepsilon u)(x) \) is \( C^\infty \) in \( R_2 \) and if \( h \) is of order \( i \) in \( R_1 \),

then

\[
D^{(s)} (J_\varepsilon u)(x) = (J_\varepsilon D^i u)(x) \text{ for } |s| \leq i
\]

in \( R_2 \).

**Proof of (iii):** We have, for each derivation \( D^{(s)} \),

\[
(D^{(s)}_x J_\varepsilon u)(x) = \int_{R_1} D^{(s)}_x j_\varepsilon(x - y) \, u(y) \, dy.
\]

Suppose \( u \) is of order \( i \) in \( R \). We have then, for \( |s| \leq i \), by partial integration,

\[
\int_{R_1} D^{(s)}_x j_\varepsilon(x - y) \, u(y) \, dy = \int_{R_1} (-1)^{|s|} [D^{(s)}_y j_\varepsilon(x - y)] \, u(y) \, dy
\]

\[
+ \int_{R_1} j_\varepsilon(x - y) \, D^{(s)} u(y) \, dy
\]
since, for each \( x \in R_2, j_\epsilon(x - y) \) considered as a function of \( y \), has compact support in \( R_1 \).

**Proof of (ii):** We have, for \( x \in R_2, \int_{R_1} j_\epsilon(x - y) \, dy = 1 \). Hence

\[
(J_\epsilon \, u)(x) - u(x) = \int_{R_1} j_\epsilon(x - y) \, (u(y) - u(x)) \, dy.
\]

By Schwarz’s inequality

\[
\int_{R_2} |(J_\epsilon \, u)(x) - u(x)|^2 \, dx \\
\leq \int_{R_2} dx \left[ \int_{R_1} j_\epsilon(x - y) \, dy \int_{R_1} j_\epsilon(x - y) \, |u(y) - u(x)|^2 \, dy \right] \\
= \int_{R_2} dx \int_{R_1} j_\epsilon(x - y) \, |u(y) - u(x)|^2 \, dy \\
\leq \int_{R_2} dx \int_{|z|<\epsilon} j_\epsilon(z) \, |u(x - z) - u(x)|^2 \, dz \\
= \int_{|z|<\epsilon} j_\epsilon(z) \left\{ \int_{R_2} |u(x - z) - u(x)|^2 \, dx \right\} \, dz
\]

Since \( \int_{R_2} |u(x - z) - u(x)|^2 \, dx \) tends to zero as \( |z| \to 0 \), (ii) is proved.

**Proof of (i):** We have, by calculations similar to the above calculations,

\[
\| J_\epsilon \, u \|_{L^2(R_2)}^2 = \int_{R_2} dx \int_{R_1} j_\epsilon(x - y) \, |u(y)|^2 \, dy \\
\leq \int_{|z|<\epsilon} j_\epsilon(z) \left\{ \int_{R_2} |u(x - z)|^2 \, dx \right\} \, dz \\
\leq \int_{|z|<\epsilon} j_\epsilon(z) \left\{ \int_{R_2} |u(x)|^2 \, dx \right\} \, dz
\]
2. Proof of Lemma 1

\[ = \| u \|_{\alpha,R_1}^2. \]

**Proof of Lemma 1.** Let \( u \) be of order in \( R_1 \) and let \( \tilde{D}^{(s)} u \) be of order \( j \) in \( R_1 \) for each \( s \) with \( |s| \leq i \). Then for \( |t| \leq j \),

\[ D^{(t)} D^{(s)} J_{\epsilon} u = D^{(t)} J_{\epsilon} \tilde{D}^{(s)} u = J_{\epsilon} \tilde{D}^{(t)} \tilde{D}^{(s)} u (|s| \leq i) \]

by (iii). Hence by (ii), \( u \) is of order \( i + j \), in \( R_1 \).

Next let \( u \) be of order \( i + j \) in \( R_1 \). Since

\[ D^{(t)} J_{\epsilon} \tilde{D}^{(s)} u = J_{\epsilon} D^{(t)} \tilde{D}^{(s)} u (|t| \leq j, |s| \leq i) \]

we see by (ii) that \( \tilde{D}^{(s)} u \) is of order \( j \) in \( R_1 \).
Part IV

Application of the semi-group theory to the Cauchy problem for the diffusion and wave equations
1 Cauchy problem for the diffusion equation

Let $R$ be a connected $n$-dimensional oriented Riemannian space with the metric

$$ds^2 = g_{ij}(x)\; dx^i \; dx^j.$$ 

Let $A$ be a second order linear partial differential operator in $R$ with $C^\infty$ coefficients:

$$(Af)(x) = b^{ij}(x) \frac{\partial^2 f}{\partial x^i \partial x^j} + a^j(x) \frac{\partial f}{\partial x^j}(x)c(x)\; f(x);$$

we assume that $b^{ij}$ is a symmetric contravariant tensor and $a^j(x)$ satisfies the transformation rule

$$a^i = a^k \frac{\partial \bar{x}_i}{\partial x_k} + b^{kl} \frac{\partial^2 \bar{x}_i}{\partial x^k \partial x^l}$$

$[(x_1, \ldots, x_n) \to (\bar{x}_1, \ldots, \bar{x}_n)]$ so that the value $(Af)(x)$ is determined independent of the choice of the local coordinates. We further assume that $A$ is elliptic in the strong sense that there exist positive constants $\mu$ and $\lambda(0 < \lambda < \mu)$ such that

$$\mu \; g^{ij}(x)\; \xi_i \; \xi_j \geq b^{ij}(x)\; \xi_i \; \xi_j \geq \lambda g^{ij}(x)\; \xi_i \; \xi_j$$

for every real vector $(\xi_1, \ldots, \xi_n)$ and every $x \in R$.

We consider the Cauchy problem in the large on $R$ for the diffusion equation: to find $u(t, x)\; (x \in R)$ such that

$$\left\{ \begin{array}{ll}
\frac{du}{dt} = A\; u(t, x), & t > 0 \\
u(0, x) = f(x), & f(x) \text{ being a given function on } R.
\end{array} \right.$$
We shall first give a rough sketch of our method of integration. We wish to integrate the equation in a certain function space $L^p(R)$ which is a Banach space (i.e., we want to obtain $u(t, \ldots) \in L^p(R)$ for each $t \geq 0$; we assume that $L^p(R)$ contains $C^\infty(R)$, the space of $C^\infty$ functions with compact support, as a dense subspace. (Examples: $L^p(R), 1 \leq p < \infty; C(R)$ if $R$ is compact). We determine an additive operator $A_o$ such that:

(i) $C^\infty(R) \supset D(A_o) \supset D^\infty(R)$, if $f \in D(A_o)A_o f = Af$.

(ii) the smallest closed extension $\bar{A}_o$ of $A_o$ exists and $\bar{A}_o$ is the infinitesimal generator of a semi group $T_t$ on $L(R)$. We then have:

$$\begin{cases}
D_t T_t f = s - \lim_{h \to 0} \frac{T_{t+h}T_t f - T_h}{h} = \bar{A}_o T_t f (= T_t \bar{A}_o f), t \geq 0 \\
T_o f = f.
\end{cases}$$

Thus $T_t f$ is a kind of solution of ($\ast\ast$). Next, we shall show that, if the initial function $f(x)$ is prescribed suitably [e.g., if $f \in D^\infty(R)$ or more generally, if $A_k f \in D(A_o)$ for all integers $k \geq 0$], there exists a function $u(t, x)$ definitely differentiable in $t$ and $x$ such that $T_t f(x) = u(t, x)$ almost everywhere in $(0, \infty) \times R$, the measure in $R$ being the one given by $\sqrt{g} d x_1, \ldots, d x_n$, and $u(t, x)$ will be a solution of ($\ast\ast$).

In carrying out this procedure, we have to solve an equation of the form $(u - \frac{A_o u}{m}) f = f$ is given and $u$ is to be found from $D(A_o)$. This is a kind of boundary value problem connected with the elliptic differential operator $A$.

**Theorem.** If $R$ is compact, the equation

$$\begin{cases}
\frac{\partial u}{\partial t} = Au = b^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + a^i(x)\frac{\partial u}{\partial x_i}, t > 0 \\
u(0, x) = f(x) \in D^\infty(R), (f(x) \text{ given})
\end{cases}$$

admits of a solution $C^\infty$ in $(t, x)$. This solution can be represented in the form

$$u(t, x) = \int_R P(t, x, dy)f(y)$$

where $P(t, x, E)$ is the transition probability of a Markoff process on $R$. 
1. Cauchy problem for the diffusion equation

The proof will be preceded by two lemmas.

We take for \( L(R) \) the Banach space \( C(R) \) of continuous functions with \( \|f\| = \sup_x |f(x)|. \mathcal{D}'(R) \) is dense in \( L(R) \). The operator \( A_o \) is defined as follows:

\[
\mathcal{D}(A_o) = \mathcal{D}'(R) \quad \text{and} \quad A_o f = Af \quad \text{for} \quad f \in \mathcal{D}'(R).
\]

**Lemma 1.** For any \( f \in \mathcal{D}'(R) \) and any \( m > 0 \), we have

\[
\max_x h(x) \geq f(x) \geq \min_{x \in R} h(x)
\]

where

\[
h(x) = f(x) - \frac{(A_o f)(x)}{m}.
\]

**Proof.** Let \( f(x) \) attain its maximum at \( x_0 \). We choose a local coordinate system at \( x_0 \) such that \( b^{ij}(x_0) = \delta_{ij} \) (Kronecker delta). □

(Such a choice is possible owing to the positive definiteness of \( b^{ij} \xi_i \xi_j \)).

Then

\[
h(x_0) = f(x_0) - m^{-1}(A_o f)(x_0)
\]

\[
= f(x_0) - m^{-1}a^i(x_0) \frac{\partial f}{\partial x^i} - m^{-1} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial (x^i)^2}
\]

\[
= f(x_0)
\]

since we have, at the maximum point \( x_0 \),

\[
\frac{\partial f}{\partial x^i} = 0 \quad \text{and} \quad \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x^i} \leq 0.
\]

Thus \( \max_x h(x) \geq f(x) \). Similarly we have \( f(x) \geq \min_x h(x) \).

**Corollary.** The inverse \( (I - m^{-1}A_o)^{-1} \) exists for \( m > 0 \) and \( \| (I - m^{-1}A_o)^{-1} \| \leq 1 \). Further \( (I - m^{-1}A_o)^{-1} h(x) \geq 0 \) if \( h(x) \geq 0 \). Also

\[
(I - m^{-1}A_o)^{-1} 1 = 1.
\]

**Lemma 2.** The smallest closed extension \( \bar{A}_o \) of \( A_o \) exists.
\( \tilde{A}_\circ f \) is defined and equal to \( h \) if there exists a sequence \( \{f_k\} \subset \mathcal{D}^\infty(R) \) such that \( s - \lim_{k \to \infty} f_k = f \) and \( s - \lim_{k \to \infty} A_\circ f_k = h \).

\( \tilde{A}_\circ f \) is determined uniquely by \( f \). For if \( \{f_k\} \subset \mathcal{D}^\infty(R) \) be such that \( \lim_{k \to \infty} f_k = 0 \) and \( \lim_{k \to \infty} A_\circ f_k = h \), then we must \( h = 0 \).

For by Green’s integral theorem, \( R \) being compact,

\[
\int_R f_k A^* g dx = \int_R g A f_k dx,
\]

for every \( g \in \mathcal{D}^\infty(R) \) so that, in the limit,

\[
0 = \int_R gh dx, \text{ for every } g \in \mathcal{D}^\infty(R); \text{ so } h = 0.
\]

To prove that the resolvent \( (I - m^{-1} A_\circ)^{-1} \) exists as a linear operator in \( C(R) \), for \( m \) large, it will be sufficient to show, in view of the Corollary to Lemma \[\text{and the fact that } \tilde{A}_\circ \text{ is closed, that the range of } (I - m^{-1} A_\circ) \text{ is dense in } C(R). \text{ We shall show that for any } h \in \mathcal{D}^\infty(R) \text{ we can find } f \in \mathcal{D}(R) \text{ such that } (I - m^{-1} A_\circ) f = h (m \text{ large}). \text{ To this purpose, we need}

\section*{2 Garding’s inequality}

For \( u, v \in \mathcal{D}^\infty(R) \), define

\[
(u, v)_0 = \int_R uv dx \quad (||u||^2_0 = (u, u)_0)
\]

\[
(u, v)_1 = (u, v)_0 + \int_R g^{ij} \frac{\partial u}{\partial x^i} \frac{\partial v}{\partial x^j} dx \quad (||u||^2_1 = (u, u)_1)
\]

Then there exists \( \gamma > 0 \) and \( \delta > 0 \) such that for all sufficiently large \( m > 0 \),

\[
B'(u, v) = \left( \left( I - \frac{A_\circ}{m} \right) u, v \right)
\]
2. Garding’s inequality

satisfies
\[ |B'(u, v)| \leq \gamma \|u\|_1 \|v\|_1 \]
\[ \delta \|u\|_1^2 \leq B'(u, u) \quad \text{for all } u, v \in \mathcal{D}^{\infty}(R). \]

This lemma can be proved by partial integration.

Let \( H_0 \) be the Hilbert space of square summable functions in \( R \). We have \( \mathcal{D}^{\infty}(R) \subset H_0(R) \). Let \( A_1 \) be the operator in \( H_0 \) with domain \( \mathcal{D}^{\infty}(R) \) defined by: \( A_1 f = A_s f, f \in \mathcal{D}^{\infty}(R) \). As in Lemma \( 2 \), the closure of \( A_1 \) in \( H_0, \tilde{A}_1 \), exists. We show now that the range of \( (I - \frac{A_s}{m}) \) is dense in \( H_0 \), for \( m \) large. If \( (I - \frac{A_s}{m}) \mathcal{D}^{\infty} \) were not dense in \( H_0 \), there will exists an element \( f \neq 0 \) in \( H_0 \) which will be orthogonal to \( (I - \frac{A_s}{m}) \mathcal{D}^{\infty} \). This mean that \( f \) is a weak solution of \( (I - \frac{A_s}{m}) f = 0 \).

By the Weyl-Schwartz theorem, \( f \) may be considered to be in \( \mathcal{D}^{\infty}(R) \). By Garding’s inequality, assuming \( m \) to be sufficiently large,
\[ \delta \|f\|_1^2 \leq \left( I - \frac{A_s}{m}, f, f \right) = 0. \text{ So } f = 0. \]

Since the range of \( (I - \frac{A_s}{m}) \) is everywhere dense in \( H_0 \), \( (I - \frac{A_s}{m})^{-1} \) is defined everywhere in \( H_0 \). So for every \( h \in \mathcal{D}^{\infty}(R) \), we can find \( f_0 \in H_0 \) such that \( f_0 \) is a weak solution of \( (I - \frac{A_s}{m}) f = h \).

Again by the Weyl-Schwartz theorem, \( f \) will be in \( \mathcal{D}^{\infty}(R) \). Thus we see that for large \( m \) the resolvent \( J_m = \left( I - \frac{A_s}{m} \right)^{-1} \) exists as a linear operator on \( L(R) \) and satisfies \( \|J_m\| \leq 1 \) (also, \( (J_m h)(x) \geq 0 \) if \( h(x) \geq 0 \); \( J_m, 1 = 1 \)). Consequently, (see Lecture 8) \( \tilde{A}_s \) is the infinitesimal generator of a uniquely determined semi-group \( T_t = \exp(t\tilde{A}_s) = s - \lim_{m \to \infty} \exp(tm(J_m - I)) \).

We have further
\[ \|T_t\| \leq 1, \quad (T_t f)(x) \geq 0 \text{ if } f(x) \geq 0, \quad T_t, 1 = 1. \]
If \( f \in \mathcal{D}^\infty(R) \), we have
\[
D_T^1 f = A \circ T f = T A f = T A_1 f \\
D_T^2 f = A \circ T A f = T A_2 f \\
\vdots \\
D_T^k f = T A_k f,
\]
for \( k \geq 0 \), since \( A_1 f \in \mathcal{D}^\infty(R) \) for integral \( k \geq 0 \). By making use of the strong continuity of \( T \) in \( t \) we see that \( (D_T^2 + \tilde{A})^k T f \) is locally square summable on the product space \((0, \infty) \times R\). Since \( (\frac{\partial^2}{\partial t^2} + A)^k \) is an elliptic operator, it follows the Friedrichs-Lax-Nirenberg theorem that \( (T f)(x) \) is almost everywhere equal to a function \( u(t,x) \) indefinitely differentiable in \( (t,x) \) for \( t \geq 0 \).

**Proof of the latter part of the theorem:**

\[
|u(t,x)| = |(T f)(x)| \leq ||T f|| \leq ||f||
\]

Hence \( u(t,x) \) is, for fixed \((t,x)\) a linear functional of \( f \in L(R) \). Therefore there exists \( P(t,x,E) \) such that
\[
u(t,x) = \int_R P(t,x,dy) f(y).
\]

The non-negativity of \( u(t,x) \) for \( f(x) \geq 0 \) implies that \( P(t,x,E) \) is \( \geq 0 \). Since \( T_i 1 = 1 \), we must have \( P(t,x,R) = 1 \).
Lecture 19

1 The Cauchy problem for the wave equation

We consider the Cauchy problem for the ‘wave equation’ in the \(m\)-dimensional Euclidean space \(E^m\):

\[
\begin{cases}
\frac{\partial^2 u(t, x)}{\partial t^2} = Au(t, x), & x \in E^m \\
u(0, x) = f(x), & u_t(0, x) g(x) \text{, } f, g, \text{ given},
\end{cases}
\]

where

\[A = d^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + b^j(x) \frac{\partial}{\partial x_j} + c(x)\]

is a second-order elliptic operator. This problem is equivalent to the matricial equation

\[
\begin{bmatrix}
\frac{\partial}{\partial t} u(t, x) \\
\frac{\partial}{\partial t} v(t, x)
\end{bmatrix} =
\begin{bmatrix}
0 & I \\
A & 0
\end{bmatrix}
\begin{bmatrix}
u(t, x) \\
v(t, x)
\end{bmatrix}
\quad (I = \text{identity}).
\]

We may apply the semi-group theory to integrate (1), by considering, in a suitable Banach-space the “resolvent equation”

\[
\begin{bmatrix}
I & o \\
o & I
\end{bmatrix} - n^{-1} \begin{bmatrix}
o & I \\
A & o
\end{bmatrix}
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix} f \\
g
\end{bmatrix}
\]

121
for large \(|n|\) \((n, \text{integral})\) and obtaining the estimate

\[
\|\left( \begin{array}{c} u \\ v \\ f \\ g \\ \end{array} \right) \| \leq (1 + |n|^{-1}|\beta|) \|\left( \begin{array}{c} f \\ g \\ \end{array} \right) \|
\]

with a positive \(\beta\) independent of \(u, f\) and \(g\). As a matter of fact, the estimate implies (see Lecture 9) that \(\begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix}\) is the infinitesimal generator of a group \(\{T_t\}_{-\infty < t < \infty}\) and

\[
T_t \begin{pmatrix} f(x) \\ g(x) \end{pmatrix} = \begin{pmatrix} u(t, x) \\ v(t, x) \end{pmatrix}
\]

will give a solution of (1) if the initial functions \(f(x)\) and \(g(x)\) are prescribed properly.

We have the

**Theorem.** Suppose that the coefficients \(a^{ij}(x), b^i(x)\) and \(c(x)\) are \(C^\infty\) and that there exists a positive constant \(\lambda\) such that

\[
a^{ij}(x)\xi_i\xi_j \geq \lambda \sum_i \xi_i^2
\]

\((x \in E^m, (\xi_1, \ldots, \xi_m) \in E^m)\). Assume further that

\[
\eta = \max \left\{ \sup_{x,i,j} |a^{ij}(x)|, \sup_{x,i,j,k} \left| \frac{\partial a^{ij}}{\partial x_k} \right|, \sup_{x,i,j,k,s} \left| \frac{\partial^2 a^{ij}}{\partial x_k \partial x_s} \right|, \sup_{x,i,j} |b^i(x)|, \sup_{x,i,j,k} \frac{\partial b^i}{\partial x_k}, \sup_x |c(x)| \right\}
\]

is finite. Then there exists a positive constant \(\beta\) such that for sufficiently small \(\alpha_*=\)

the equation (1) is solvable in the following sense: for any pair of \(C^\infty\) functions \([f(x), g(x)]\) on \(E^m\) for which \(A_k f, A_k g\) and their partial derivatives are square integrable (for each integer \(k \geq 0\)) over \(E^m\), the equation (1) admits of a \(C^\infty\) solution \(u(t, x)\) satisfying the “energy inequality”

\[
\left( (u - \alpha_0Au, u) + \alpha_0(u, u) \right)^{\frac{1}{2}} \leq \exp(\beta|t|((f - \alpha_0Af, f) + \alpha_0(g, g)))^{\frac{1}{2}}
\]
The proof will be carried out in several steps.

**First step:** Let $H$ be the space of real-valued $C^\infty$ functions in $\mathbb{E}^m$ for which

$$||f||_1 = \left\{ \int_{\mathbb{E}^m} |f|^2 \, dx + \sum_i \int_{\mathbb{E}^m} |f_i|^2 \, dx \right\}^{\frac{1}{2}} < \infty,$$

and let $\tilde{H}_1(\mathbb{E}^m)$ be the completion of $H$ with respect to the norm $|| ||_1$. The completion of $H$ with respect to $||f||_0 = \left\{ \int_{\mathbb{E}^m} |f|^2 \, dx \right\}^{\frac{1}{2}}$ will be denoted by $\tilde{H}_0(\mathbb{E}^m)$. $\tilde{H}_0(\mathbb{E}^m)$ and $\tilde{H}_1(\mathbb{E}^m)$ are Hilbert spaces; actually $\tilde{H}_0(\mathbb{E}^m) = H_0(\mathbb{E}^m) = L_2(\mathbb{E}^m)$.

One can prove that there exists $\chi > 0$ and $\alpha_0 > 0$ such that for $0 < \alpha < \alpha_0$ there correspond $\gamma > 0$ and $\delta_\alpha > 0$ satisfying

$$\delta_\alpha ||f||_1^2 \leq \begin{cases} (f - \alpha Af, f)_o & \text{for } f \in H, Af \in H_0 \\ (f - \alpha A^* f, f)_o & \text{for } f \in H, A^* f \in H_0. \end{cases}$$

$$||f - \alpha Af, g||_o \leq (1 + \alpha \gamma) ||f||_1 ||g||_1 \text{ for } f, g \in H, Af \in H_0.$$

$$||f - \alpha A^* f, g||_o \leq (1 + \alpha \gamma) ||f||_1 ||g||_1 \text{ for } f, g \in H, A^* f \in H_0.$$

$$|(af, g)_o - (f, Ag)_o| \leq \chi ||f||_1 ||g||_o \text{ for } f, g \in H, Af, Ag \in H_0.$$

(The proofs of these inequalities will be given in the next lecture). □

Thus the bilinear form

$$B_\alpha^o(u, v) = (u - \alpha A^* u, v)_o \text{ for } u, v \in H, A^* u \in H_0$$

can be extended to a bilinear functional $B_\alpha(u, v)$ on $H_1$ satisfying

$$\begin{cases} \delta_\alpha ||u||_1^2 \leq B_\alpha(u, u) \\ |B_\alpha(u, v)| \leq (1 + \alpha \gamma) ||u||_1 ||v||_1. \end{cases}$$

**Second step:** Let $0 < \alpha \leq \alpha_0$. For any $f \in H$, the equation $u - \alpha Au = f$ admits of a uniquely determined solution $u(x) = u_f(x) \in H$. 

1. The Cauchy problem for the wave equation

Proof. The proof will be carried out in several steps.
Proof. The additive functional $F(u) = (u, f)_o$ is bounded on $H_1$, because
\[
|F(u)| = |(u, f)_o| \leq \|u\|_o \|f\|_o \leq \|u\|_1 \|f\|_o.
\]
So, by Riesz’s representation theorem, there exists a uniquely determined $v(f) \in \tilde{H}_1$ such that
\[
(u, f)_o = (u, v(f))_1.
\]

By the Lax-Milgram theorem (see lecture 4) as applied to the bilinear form $B_\alpha(u, v)$ in $\tilde{H}_1$, there corresponds a uniquely determined element $S v(f)$ in $\tilde{H}_1$ such that
\[
(u, f)_o = (u, v(f))_1 = B_\alpha(u, S v(f)), \quad \text{for } u \in H_1.
\]

$u_o = S v(f)$ is a weak solution of the equation $u - \alpha Au = f$, i.e., for each $u \in \mathcal{D}^\infty(R)$ we have $(u, f)_o = (u - \alpha A^* u, S v(f))_o$. In fact, let $\{v_k\} \subset H$ be a sequence such that $v_k \rightarrow S v(f)$ in $\tilde{H}_1$; then, for
\[
\begin{align*}
\lim_{n \rightarrow \infty} B_\alpha(u, v_n) &= \lim_{n \rightarrow \infty} (u - \alpha A^* u, v_n) \\
&= (u - \alpha A^* u, S v(f)).
\end{align*}
\]

Since $f$ is $C^\infty$ in $E^m$ and $A$ is elliptic, $u_o = S v(f)$ is almost everywhere equal to a $C^\infty$ function (Weyl-Schwartz theorem). We thus have a solution $u_o \in H$ of the equation $u - \alpha Au = f$. The uniqueness of the solution follows from the inequalities given in the first step. □

Third step: If the integer $n$ is such that $|n|^{-1}$ is sufficiently small, then for any pair of functions $\{f, g\}$ with $f, g \in H$ and $Af \in H_o$, the equation
\[
\begin{pmatrix}
I & 0 \\
0 & I
\end{pmatrix}
- n^{-1}
\begin{pmatrix}
I & 0 \\
A & 0
\end{pmatrix}
\begin{pmatrix}
u \\
v
\end{pmatrix} = \begin{pmatrix} f \\
g
\end{pmatrix}
\]

or
\[
u - n^{-1}v = f
\]
admits of a uniquely determined solution \( \{u, v\} \) \( u, v \in H \). Moreover, we have

\[
[B_{\alpha,\alpha}(u, u) + \alpha_{\alpha}(v, v)]^{\frac{1}{2}} \leq \left( 1 + |n|^{-1} \right) \beta (B_{\alpha,\alpha}(f, f) + \alpha_{\alpha}(g, g))^{\frac{1}{2}}
\]

with a positive constant \( \beta \).

**Proof.** Let \( u_1, v_1 \in H \) be such that

\[
u_1 - n^{-2} Au_1 = f \quad v_2 - n^{-2} Av_2 = g.
\]

(See the second step). Then

\[
u = u_1 + n^{-1} v_2 \quad v = n^{-1} Au_1 + v_2
\]

satisfies (2).

We have

\[
Au = n(v - g) \in H \subset H_o, \quad Av = n(Au - Af) \in H_o.
\]

We may therefore apply the inequalities of the first step.

Thus by (2),

\[
(f - \alpha_{\alpha}Af, f)_o = (u - n^{-1}v - \alpha_{\alpha}(u - n^{-1}v), u - n^{-1}v)_o
\]

\[
= (u - \alpha_{\alpha}Au, u)_o - 2n^{-1}(u, v)_o + \alpha_{\alpha}n^{-1}(Au, v)_o
\]

\[
+ \alpha_{\alpha}n^{-2}(Av, u)_o + n^{-2}(v - \alpha_{\alpha}Av, v)_o
\]

and

\[
\alpha_{\alpha}(g, g)_o = \alpha_{\alpha}(v - n^{-1}Au, v - n^{-1}Au)_o
\]

\[
= \alpha_{\alpha}(v, v)_o - \alpha_{\alpha}n^{-1}(v, Au)_o - \alpha_{\alpha}n^{-1}(Au, v)_o + \alpha_{\alpha}n^{-2}(AA)_o
\]

Hence

\[
B_{\alpha,\alpha}(f, f)_o + \alpha_{\alpha}(g, g) \geq B_{\alpha,\alpha}(u, u) + \alpha_{\alpha}(v, v)_o - 2|n|^{-1}(u, v)_o - \alpha_{\alpha}|n|^{-1}(Av, u)_o - (Au, v)_o
\]
\[ B\alpha_{\circ}(u, u) + \alpha_{\circ}(u, v) - 2|n|^{-1} \|u\|_1 \|v\|_\infty - \alpha_{\circ}|n|^{-1} \chi |\|u\|_1 \|v\|_\infty \].

Thus, by taking \( \tau > 0 \) sufficiently large and then taking \(|n|\) sufficiently large, we have the desired inequality. \( \square \)

Fourth step: The product space \( \tilde{H}_1 \times \tilde{H}_\circ \) is a Banach space with the norm

\[ \|(u)\| = [B\alpha_{\circ}(u, u) + \alpha_{\circ}(v, v)]^{1/2}. \]

We define now an operator \( \mathcal{O} \) in \( \tilde{H}_1 \times \tilde{H}_\circ \): the domain of \( \mathcal{O} \) consists of the vectors \( (u, v) \in H_1 \times H_\circ \) such that \( u, v \in H \) and \( A(u - n^{-1}v) \in H_\circ \) and \( v - n^{-1}Au \in H \) and on such elements \( \mathcal{O}(u, v) \) is defined to be

\[ \begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \]

The third step shows that for sufficiently large \(|n|\), the range of the operator \( \mathcal{O} \) coincides with the set pairs \( (f, g) \) such that \( f, g \in H, Af \in H_\circ \); such vectors \( (f, g) \) are dense in the Banach space \( \tilde{H}_1 \times \tilde{H}_\circ \). It follows that the smallest closed extension \( \tilde{\mathcal{O}} \) of \( \mathcal{O} \) admits, for sufficiently large \(|n|\), of an inverse \( \mathcal{J}_n = (\mathcal{J} - n^{-1} \tilde{\mathcal{O}})^{-1} \) which is linear operator on \( \tilde{H}_1 \times H_\circ \) satisfying

\[ \|\mathcal{J}_n\| \leq (1 + \beta |n^{-1}|). \]

So, there exists a uniquely determined group \( \{T_t\}_{t \in (-\infty, \infty]} \) with \( \tilde{\mathcal{O}} \) as the infinitesimal generator and such that

\[ \|T_t\| \leq \exp(\beta t), \]

\[ \lim_{h \to 0} \frac{T_{t+h} - T_t}{h} = \tilde{\mathcal{O}} T_t (f, g) = T_t \mathcal{O} (f, g) \quad \text{if} \quad (f, g) \in \text{domain of} \ \mathcal{O} \quad \text{(See Lecture 9)}. \]
Lecture 20

1 Cauchy problem for the wave equation (continued)

Fifth step: If \( f \) and \( g \) satisfy the conditions of the theorem, i.e., if \( A^k f \in H, A^k g \in H(k = 0, 1, \ldots) \), we have

\[
\bar{\partial}^k \begin{pmatrix} f \\ g \end{pmatrix} = \bar{\partial}^k \begin{pmatrix} f \\ g \end{pmatrix} \in \bar{H}_1 \times H_{\sigma}(k = 0, 1, \ldots),
\]

i.e., \( \begin{pmatrix} f \\ g \end{pmatrix} \) is in the domain of every power of \( \bar{\partial}^k \). So, referring to step 4, we find that vectors

\[
\begin{pmatrix} v(t, x) \\ v(t, x) \end{pmatrix} = T_t \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}
\]

are in the domain of every power of \( \bar{\partial} \) :

\[
\bar{\partial}^k \begin{pmatrix} u(t, x) \\ v(t, x) \end{pmatrix} \in \bar{H}_1 \times H_{\sigma}(k = 0, 1, 2, \ldots)
\]

Thus, for integral \( k \geq 0, u(t, x) \) is for \( t \) fixed, a weak solution of the equation

\[
A^k u = f^{(k)}, \text{ with } f^{(k)} \in L_2(E^m)
\]

\( A^k \) is an elliptic operator of order \( 2k \) and \( k \) may be taken arbitrarily large. We see therefore by the Friedrichs-Lax-Nirenberg theorem and Sobolev’s lemma, that \( u(t, x) \) is \( C^{\infty} \) in \( x \) (for fixed \( t \)). And the same statement holds for \( v(t, x) \).
Since $\|T_t\| \exp \beta(|t|)$ we see that
\[ \|u(t,x)\|_1^2 + \|v(t,x)\|_1^2 \leq \text{Const.} \exp(2\beta(|t|)} \left( \|f\|_1^2 + \|g\|_0^2 \right) . \]

This, combined with the strong continuity of $T_t$ in $t$, shows that $u(t,x)$ and $v(t,x)$ are locally square summable in the product space $(-\infty < t < \infty) \times E^m$. And we have, for the second order strong derivative $\partial_t^2$,
\[ \partial_t^2 u(t,x) = Au(t,x) \]
so that $(\partial_t^2 + A)u = 2Au, (\partial_t^2 + A)^ku = (2A)^ku$.

Since $\partial_t^2 + A$ is an elliptic operator in $(-\infty < t < \infty) \times E^m$, we see that $u(t,x)$ is almost everywhere equal to a function $C^\infty$ in $(t,x)$. The proof of the first step is obtained by the Lemma.

Let $f, g \in H$ and $Af \in H_o$. Then
\[ (Af, g)_o = -\int \delta^{ij} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} dx - \int b^{ij} \frac{\partial f}{\partial x_i} g dx + \int c f g. \]

And we can also partially integrate the second and the third terms on the right, so that the first order derivatives of $\frac{\partial f}{\partial x_i}$ shall be eliminated, and the integrated terms are nought.

**Proof.** From $Af \in H_o$ and $g \in H$ we see that $\delta^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} g$ is integrable over $E^m$. Thus, by Fubini theorem,
\[
\int_{E^m} \delta^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} g dx = \lim_{\delta \to 0} \int_{-\infty}^{\infty} \cdots dx_m \int_{x_1}^{\delta_1} \cdots \int_{x_1}^{\delta_1} \delta^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} g dx_1 \cdots dx_1
\]
\[
\int_{E_1} \delta^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} g dx_1 \bigg|_{x_1 = \delta_1} - \int \delta^{ij} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} dx_1
\]
\[
+ \int \delta^{ij} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} dx_1 + \int \sum_{i \neq j} \delta^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} g dx_1. \]
1. Cauchy problem for the wave equation (continued)

\[ = k_1(\delta_1, \varepsilon_1, x_2, \ldots, x_m) + k_2(\delta_1, \varepsilon_2, x_2, \ldots, x_m) + k_3(\delta_1, \varepsilon_1, x_2, \ldots, x_m) \]

By Schwarz’s inequality, we have

\[ \left| \int_{-\infty}^{\infty} dx_2 \ldots dx_m k_1 \right| \]

\[ \leq \eta \sum_{j} \int_{-\infty}^{\infty} dx_2 \ldots dx_m \left| \frac{\partial f(\delta_1, x_2, \ldots, x_m)}{\partial x_j} \right|^2 \int_{-\infty}^{\infty} dx_2 \ldots dx_m |g(\delta_1, x_2, \ldots, x_m)|^2 \]

+ similar terms pertaining to \( \varepsilon_1 \) instead of \( \delta_1 \).

Since

\[ \int_{E^n} g^2 dx = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_m)^2 dx_2, \ldots, dx_m, \]

\[ \int_{E^n} \left| \frac{\partial f}{\partial x_j} \right|^2 dx = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} \| \frac{\partial f}{\partial x_j}(x_1, x_2, \ldots, x_m) \|^2 dx_2 \ldots dx_m, \]

we see there exists \( \{ \delta_1^{(n)} \} \) and \( \{ \varepsilon_1^{(n)} \} \) such that

\[ \lim_{\delta_1^{(n)} \to \infty} \int k_1(\varepsilon_1^{(n)}, \delta_1^{(n)}, x_2, \ldots, x_m) dx_2 \ldots dx_m = 0. \]

On the other hand, since \( f, g, \frac{\partial f}{\partial x_j}, \frac{\partial g}{\partial x_1} \in H_o \), we see that

\[ \lim_{\delta_1 \to \infty} \int k_2(\varepsilon_1, \delta_1, x_2, \ldots, x_m) dx_2, \ldots, dx_m \]

\[ = \int_{E^n} \left\{ -a_{ij} \frac{\partial f}{\partial x_j} \frac{\partial g}{\partial x_1} - \frac{\partial^2 f}{\partial x_1 \partial x_j} \right\} dx = k_2 \]

is finite. Thus,

\[ \int a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} g dx = k_2 + \lim_{\delta_1^{(n)} \to \infty} \int_{-\infty}^{\infty} k_3(\varepsilon_1^{(n)}, \delta_1^{(n)}, x, x_2, \ldots, x_m) dx_2 \ldots dx_m \]
Hence
\[
\int_{\varepsilon_1^{(n)}}^\delta \sum_{i,j \neq 1} a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \, gd\!x_1
\]
is integrable over \(-\infty < x_1 < \infty(i = 2, \ldots, m)\). □

Hence
\[
k_3 = \lim_{\delta_1^{(n)} \to \infty} \lim_{\varepsilon_1^{(n)} \to -\infty} \int_{-\infty}^{\infty} dx_2 \ldots dx_m k_3(\varepsilon_1^{(n)}, \delta_1^{(n)}, x_2, \ldots, x_m)
\]
\[
= \lim_{\delta_1^{(n)} \to \infty} \lim_{\varepsilon_1^{(n)} \to -\infty} \int_{-\infty}^{\infty} dx_2 \ldots dx_m \left\{ \int_{\varepsilon_2}^{\delta_2} \sum_{i,j \neq 1} a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \, gd\!x_1 \right\}
\]

However
\[
\{ \ldots \} = \int_{\varepsilon_1^{(n)}}^{\delta_1^{(n)}} \int_{\varepsilon_2}^{\delta_2} \sum_{i,j \neq 1} -a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \, dx_2
\]
\[
= \int_{\varepsilon_1^{(n)}}^{\delta_1^{(n)}} \left[ a_{ij} \frac{\partial f}{\partial x_j} \right]_{x_2=\varepsilon_2}^{x_2=\delta_2} + \int_{\varepsilon_2}^{\delta_2} -a_{ij} \frac{\partial f}{\partial x_j} \frac{\partial g}{\partial x_2} \, dx_2
\]
\[
- \int_{\varepsilon_2}^{\delta_2} \frac{\partial a_{ij}}{\partial x_2} \frac{\partial f}{\partial x_j} \, gd\!x_2 + \int_{\varepsilon_2}^{\delta_2} \sum_{i,j \neq 1,2} \frac{\partial^2 f}{\partial x_i \partial x_j} gd\!x_1 dx_2
\]
we have
\[
\left| \int_{-\infty}^{\infty} dx_3 \ldots dx_m \int_{\varepsilon_1^{(n)}}^{\delta_1^{(n)}} a_{ij} \frac{\partial f}{\partial x_j} \, gd\!x_1 \right|
\]
1. Cauchy problem for the wave equation (continued)

\[ \leq \eta \sum_j \left( \int dx_1 dx_3 \ldots dx_m \left| \frac{\partial f}{\partial x_j} \right|^2 \int g^2 dx_1 dx_3 \ldots dx_m \right)^{1/2} \]

and so, by the integrability on \( E^m \) of \( \left| \frac{\partial f}{\partial x_j} \right|^2 \) and \( \left| g \right|^2 \), there exists \( \delta_{(2)}^{(1)}, \epsilon_{(2)}^{(1)} \) such that

\[ \lim_{\delta_{(2)}^{(1)} \to \infty} \int dx_3 \ldots dx_m \int_{\epsilon_1^{(1)}} \left[ a^{2j} \frac{\partial f}{\partial x_j} g \right]_{x_2 = E^{(1)}} dx_1 = 0 \]

uniformly with respect to \( \delta_1 \) and \( \epsilon_1 \).

We have also

\[ \lim_{\delta_{(1)}^{(n)} \to \infty} \int dx_3 \ldots dx_m \int_{\epsilon_1^{(n)}} \left[ a^{2j} \frac{\partial f}{\partial x_j} g \right]_{x_2 = E^{(1)}} dx_1 \]

\[ = \int_{E^m} \left( a^{2j} \frac{\partial f}{\partial x_j} g - \frac{\partial a^{2j}}{\partial x_2} \frac{\partial g}{\partial x_j} \right) dx. \]

Therefore

\[ \int a^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} g dx = - \int \sum_{i \neq j = 1, 2} a^{ij} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} dx - \int \sum_{i, j = 1, 2} \frac{\partial a^{ij}}{\partial x_i} \frac{\partial f}{\partial x_j} g dx \]

\[ + \lim_{\delta_{(1)}^{(n)} \to \infty} \lim_{\delta_{(2)}^{(n)} \to \infty} \int dx_3 \ldots dx_m \int_{\epsilon_1^{(n)}} \int_{\epsilon_2^{(n)}} \sum_{i, j = 1, 2} a^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} g dx_1 dx_2 \]

Repeating the process, we get the Lemma.
Lecture 21

1 Integration of the Fokker-Planck equation

We consider the Fokker-Planck equation

\[
\frac{\partial u(t, x)}{\partial t} = Au(t, x), \quad t \geq 0
\]

\[
(Af)(x) = \frac{1}{\sqrt{g(x)}} \frac{\partial^2}{\partial x^i \partial x^j} (\sqrt{g(x)} d^{ij}(x)f(x)) - \frac{1}{\sqrt{g(x)}} \frac{\partial}{\partial x^i} (g(x)b^i(x)f(x))
\]

in a relatively compact subdomain \( R \) (with a smooth boundary) of an oriented \( n \)-dimensional Riemannian space with the metric \( ds^2 = g_{ij}(x)dx^i dx^j \). As usual the volume element in \( R \) is defined by \( dx = \sqrt{g(x)} dx^1 \cdots dx^n \), where \( g(x) = \det(g_{ij}(x)) \). We assume that the contravariant tensor \( a^{ij}(x) \) is such that \( a^{ij}\xi_i\xi_j > 0 \) for \( \sum_{i=1}^m \xi_i^2 > 0, \xi_i \) real. The functions obey, for the coordinate transformation \( x \to \bar{x} \), the transformation rule

\[
\bar{b}^i(\bar{x}) = \frac{\partial \bar{x}}{\partial x^k} b^k + \frac{\partial^2 \bar{x}}{\partial x^k \partial x^s} a^{ks}(x).
\]

We assume that \( g_{ij}(x), a^{ij}(x) \) and \( b^i(x) \) are \( C^\infty \) function of the local coordinates \( x = (x^1 \cdots x^n) \).

Suggested by the probabilistic interpretation of the Fokker-Planck equation due to A. Kolmogorov, we shall solve the Cauchy problem in the space \( L_1(R) \).
Green’s integral formula:

Let \( A^* \) be the formal adjoint of \( A \);

\[
A^* = a^{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j} + a^i(x)\frac{\partial}{\partial x^i}.
\]

Let \( G \) be a subdomain of \( R \) with a smooth boundary \( \partial G \). Then we obtain by partial integration Green’s formula:

\[
\int_G \{ h(x)(Af)(x) - f(x)(A^*h)(x) \} dx
= \int_{\partial G} \left( \frac{\partial}{\partial x^j} \sqrt{g(x)}a^{ij}(x) - \sqrt{g(x)}b^i(x) \right) \cos(n, x^i) f(x) h(x) dS
+ \int_{\partial G} \sqrt{g(x)}a^{ij}(x) \left( h(x) \frac{\partial f}{\partial x^j} - f(x) \frac{\partial h}{\partial x^j} \right) \cos(n, x^i) dS
\]

where \( n \) denotes the outer normal at the point \( x \) of \( \partial G \) and \( ds \) denotes the hypersurface area of \( \partial G \).

**Remark.** If \( a^{ij}(x) \cos(n, x^i) \cos(n, x^j) > 0 \) at \( x \in \partial G \) we may define the transversal (or conormal) direction \( \nu \) at \( x \) by

\[
\sqrt{g(x)}a^{ij}(x) \cos(n, x^i)dS_i = dv(i = 1, 2, \ldots, m)
\]

so that we have

\[
\sqrt{g(x)}a^{ij}(x)(h(x) \frac{\partial f}{\partial x^j} - f(x) \frac{\partial h}{\partial x^j}) \cos(n, x^i) dS
= (h(x) \frac{\partial f}{\partial v} - f(x) \frac{\partial h}{\partial v}) dS.
\]

We consider \( A \) to be an additive operator defined on the totality of \( D(A) \) of \( C^\infty \) functions \( f(x) \) in \( RU\partial R \) with compact supports satisfying the following boundary condition:

\[
\sqrt{g(x)}a^{ij}(x) \frac{\partial f}{\partial x^j} \cos(n, x^i) + \left( \frac{\partial}{\partial x^j} \sqrt{g(x)}a^{ij}(x) - \sqrt{g(x)}a^i(x) \right) \cos(n, x^i) f(x) = 0.
\]
1. Integration of the Fokker-Planck equation

(When \( R \) is a subdomain of the euclidean space \( E^m \) and \( A \) is the Laplacian \( \Delta \) the above condition is nothing but the co called “reflecting barrier condition”)

\[
\frac{\partial f}{\partial n} = 0,
\]

since \( \nu \) and \( n \) coincide in this case). \( D(A) \) is dense in the Banach space \( L_1(R) \).

To discuss the resolvent of \( A \) we begin with

**Lemma 1.** Let \( f(x) \in D(A) \) be positive (or negative) in domain \( G \subseteq R \) such that \( f(x) \) vanishes on \( \partial G - \partial R \), (i.e., \( f(x) \) vanishes on the part of \( \partial G \) not contained in \( \partial R \)). Then we have the inequality

\[
\int_G (Af)(x)dx \leq 0 \quad \text{resp.} \quad \int_G Af(x)dx \geq 0.
\]

**Proof.** Taking \( h \equiv 1 \) in Green’s formula and remembering the boundary condition on \( f(x) \), we obtain

\[
\int_G (Af)(x)dx = \int_{\partial G - \partial R} \frac{\partial f}{\partial n} ds \\
\leq 0.
\]

□

**Corollary.** For \( f \in D(A) \) we have for any \( \alpha > 0 \) \( \| f - \alpha^{-1}Af \| \geq \| f \| \).

**Proof.** Let \( h(x) = 1, -1 \) or 0 according as \( f(x) \) is \( > 0, < 0 \) or \( = 0 \). Since the conjugate space of \( L_1(R) \) is the space of essentially bounded function \( k(x) \) with the norm

\[
\| k \| = \text{essential sup}_{x \in R} |k(x)|,
\]

we have

\[
\| f - \alpha^{-1}Af \| \geq \int_R h(x) \left( f(x) - \alpha^{-1}Af(x) \right) dx
\]
\[
\int |f(x)|dx - \alpha^{-1} \sum_i \int_{P_i} (Af)(x)dx
\]
\[+ \alpha^{-1} \sum_i \int_{N_j} (Af)(x)dx
\]

where \( P \) (resp. \( N \)) is connected domain in which \( f(x) > 0 \) (resp. \( < 0 \)) such that \( f(x) \) vanishes on \( \partial P \) (resp. \( \partial N \)).

**Lemma 2.** The smallest closed extension \( \tilde{A} \) of \( A \) exists and for any \( \alpha > 0 \) the operator \( (I - \alpha^{-1} \tilde{A}) \) admits of a bounded inverse, \( J_\alpha = (I - \alpha^{-1} \tilde{A})^{-1} \) with norm \( \leq 1 \).

**Proof.** The existence of \( \tilde{A} \) follows from Green’s formula. For if \( \{f_k\} \subseteq D(A) \) be such that strong \( \lim f_k = 0 \), strong \( \lim Af_k = h \), then for \( \varphi \in \mathcal{D}_\infty (\mathbb{R}) \),

\[
\lim_{R} \int_{R} \{\varphi Af_k - f_k A^* \varphi \} dx = 0, \quad \text{(or)}
\]
\[
\int_{\mathbb{R}} \varphi hd\mathbf{x} = 0. \quad \text{So } h = 0.
\]

The other part of the lemma follows from the corollary of lemma □

**Lemma 3.** \( \tilde{A} \) is the infinitesimal generator of a semi-group \( T_t \) in \( L_1(\mathbb{R}) \) if and only if for sufficiently large \( \alpha \) the range \( \{(I - \alpha^{-1} A)f, f \in D(A)\} \) of the operator \( (I - \alpha^{-1} A) \) is dense in \( L_1(\mathbb{R}) \). Moreover, if this condition is satisfied, then \( J_\alpha \) is a transition operator, i.e., if \( f(x) \geq 0 \) and \( f \in L_1(\mathbb{R}) \), then \( (J_\alpha f)(x) \geq 0 \) and

\[
\int_{\mathbb{R}} (J_\alpha f)(x)dx = \int_{\mathbb{R}} f(x)dx.
\]

**Proof.** The first part is evident. Then latter part may be proved as follows: For any \( g(x) \geq 0 \) of \( L_1(\mathbb{R}) \) there exists a sequence \( \{f_k(x)\} \subseteq D(A) \)
1. Integration of the Fokker-Planck equation

such that \( s - \lim f_k = f \) exists and \( s - \lim (f_k - \alpha^{-1}A f_k) = f - \alpha^{-1} \tilde{A} f = g \).

By the boundary condition on \( f_k \), we have

\[
\int_R (f_k - \alpha^{-1}A f_k) \, dx = \int_R f_k \, dx,
\]

(Put \( h(x) \equiv 1 \) in Green’s formula). So in the limit we have

\[
\int_R g \, dx = \int_R f \, dx.
\]

Also, by the Corollary to Lemma I

\[
\int_R |f_k - \alpha^{-1}A f_k| \, dx \geq \int_R |f_k| \, dx,
\]

and hence

\[
\int_R |g| \, dx \geq \int_R |f| \, dx.
\]

Therefore by the positivity of \( g(x) \)

\[
\int f(x) \, dx = \int g(x) \, dx = \int |g(x)| \, dx \geq \int |f(x)| \, dx
\]

proving that \( J_\alpha \) is a transition operator. \( \square \)

Therefore the semi-group

\[
T_t u = \lim_{\alpha \to \infty} \exp(t \tilde{A} J_\alpha) u = \lim_{\alpha \to \infty} \exp(\alpha t (J_\alpha - I) u)
\]

is a semi-group of transition operators.
1 Integration of the Fokker-Planck equation
(Continued)

Before going into the proof of the differentiability of the operator-theoretical solution \( u(t, x) = (T_t u)(x) \) we shall discuss the question of the denseness of the range of the set

\[
\{ (I - \alpha^{-1} A)f, f \in D(A) \}.
\]

If the range of \( (I - \alpha^{-1} A) \) were not dense in \( L_1(R) \), there will exist \( h \in M(R) = L_1(R)^*, h \neq 0 \) such that

\[
\int_{R} (I - \alpha^{-1} A)f \cdot h \, dx = 0, f \in D(A).
\]

\( h \) is a weak solution of the equation \( (I - \alpha^{-1} A^*)h = 0 \). Since \( h \in L_2(R) \) and \( A^* \) is elliptic, \( h \) is almost everywhere equal to a bounded \( C^\infty \) solution of \( (I - \alpha^{-1} A^*)h = 0 \). Let \( \{R_k\} \) be a monotone increasing sequence of domains \( \subseteq R \) with smooth boundary such that \( \partial R_k \) tends to \( \partial R \) very smoothly. Then we have

\[
0 = \int_{R_k} h(I - \alpha^{-1} A)f \, dx - \int_{R} f(I - \alpha^{-1} A^*)h \, dx
= \alpha^{-1} \int_{R_k} (hf - fA^*h) \, dx
\]
\[ = \alpha^{-1} \left\{ \int_{\partial R_k} \sqrt{g} d^j \left( h \frac{\partial f}{\partial x_j} - f \frac{\partial h}{\partial x_j} \right) \cos(n, x') dS \\
+ \int_{R_k} \left( \frac{\partial \sqrt{g} d^j}{\partial x'} - \sqrt{g} b^j \right) \cos(n, x') f(x) h(x) dS \right\}. \]

By the boundedness of \(h\) and the boundary condition on \(f\), we have

\[ \lim_{k \to \infty} \int_{\partial R_k} \left\{ \sqrt{g} d^j \frac{\partial f}{\partial x'} + \left( \frac{\partial \sqrt{g} d^j}{\partial x'} - \sqrt{g} b^j \right) f \right\} \cos(n, x') h dS = 0. \]

Therefore \(h\) must satisfy the boundary condition

\[ \lim_{k \to \infty} \int_{\partial R_k} \sqrt{g} d^j \frac{\partial h}{\partial x'} \cos(n, x') dS = 0 \text{ for every } f \in D(A). \]

Such a bounded solution \(h\) of \((I - \alpha^{-1} A^*)h = 0\) is identically zero and hence \(\bar{A}\) is the infinitesimal generator of a semi-group \(T_t\) in \(L_1(\mathbb{R})\) in either of the following cases:

(i) \(R\) is compact (without boundary).

(ii) \(R\) is a half-line or a finite closed interval on the real line and \(A = d^2/dx^2\).

**Proof.**

(i) At a maximum (or minimum) point \(x_0\) of \(h(x)\) we must have \(A^* h(x_0) \leq 0\) (resp. \(\geq 0\)) so that the continuous solution \(h(x)\) of \(A^* h = \alpha h\) cannot have either a positive maximum or a negative minimum.

(ii) The boundary condition for \(h\) is \(\frac{\partial h}{\partial n} = 0\) and the general solution of \(A^* h = \alpha h\) is

\(\square\)
1. Integration of the Fokker-Planck...

\[ h = C_1 e^{\sqrt{\alpha} x} - C_2 e^{-\sqrt{\alpha} x} \]

\[ \frac{dh}{dx} = C_1 \sqrt{\alpha} x + C_2 \sqrt{\alpha} - e^{\sqrt{\alpha} x} \]

so that the vanishing of \( \frac{dh}{dx} \) at two points implies that \( C_1 = C_2 = 0 \). And the vanishing of \( \frac{dh}{dx} \) at one point implies either \( C_1 = C_2 = 0 \) or \( C_1 \) and \( C_2 \neq 0 \). The latter contingency contradicts the boundedness of \( h \).

**A parametrix for the operator** \(-\left( \frac{\partial}{\partial \tau} + A^* \right)\)

Let \( \Gamma(P, Q) = r(P, Q)^2 \) be the square of the shortest distance between the points \( P \) and \( Q \) according to the metric \( dr^2 = a_{ij} dx^i dx^j \), where \( (a_{ij}) = (a^{ij})^{-1} \). We have the

**Theorem.** For any positive \( k \) we may construct a parametrix \( H_k(P, Q, t - \tau) \) for \(-\left( \frac{\partial}{\partial \tau} + A^* \right)\) of the form

\[ H_k(P, Q, t - \tau) = (t - \tau)^{-m/2} \exp \left( -\frac{\Gamma(P, Q)}{4(t - \tau)} \sum_{i=0}^{k} u_i(P, Q)(t - \tau)^i \right), \]

where \( u_i(P, Q) \) are \( C^\infty \) functions in a vicinity of \( P \) and \( u_i(P, P) = 1 \), we have

\[ \left( -\frac{\partial}{\partial \tau} - A^*_{\nu} \right) H_k(P, Q, t - \tau) = (t - \tau)^{k-m/2} \exp \left( -\frac{\Gamma(P, Q)}{4(t - \tau)} \right) C_k(P, Q) \]

\( C_k(P, Q) \) being \( C^\infty \) functions in a vicinity of \( P \).

**Proof.** We introduce the normal coordinates* \( y^\sigma \) of the point \( Q = (x^1, \ldots, x^m) \) in suitable neighbourhood of \( P \).

\[ y^\sigma = [\Gamma(P, Q)]^{1/2} \left( \frac{dx^\sigma}{dr} \right)_{r=Q} \]

Let \( dr^2 = \alpha_{ij}(y)dy^i dy^j \). We first show that when we apply the operator

\[ A^* = Ay^\nu = \alpha^{ij} \frac{\partial^2}{\partial y_i \partial y_j} + \beta^i \frac{\partial}{\partial y^i} + e \]
on a function $f(\Gamma, y)$ ($\Gamma$ is function on $y$) we have,

$$A^*_y = 4\Gamma \frac{\partial^2 f}{\partial \Gamma^2} + 4y^\sigma \frac{\partial^2 f}{\partial \Gamma \partial y^\sigma} + M \frac{\partial f}{\partial \Gamma} + N(f)$$

$$N(f) = a^{ij} \frac{\partial^2 f}{\partial y^i \partial y^j} + \beta^i \frac{\partial f}{\partial y^i} + ef$$

(The differentiations have to be performed as though $\Gamma$ and $y$ were independent variables). To prove this, we need the well-known formulae:

$$\Gamma(P, Q) = \alpha_{ij}(0) y^i y^j$$

$$\alpha_{ij}(y^j) = \alpha_{ij}(0) y^j.$$  

Define

$$\frac{d}{dy^i} f(y, \Gamma) = \frac{\partial f}{\partial y^i} + \frac{\partial f}{\partial \Gamma} \frac{\partial \Gamma}{\partial y^i}.$$ 

Then

$$\frac{\partial^2}{\partial y^i \partial y^j} [f(y, \Gamma)] = \frac{\partial}{\partial y^i} \left( \frac{\partial f}{\partial y^j} + \frac{\partial f}{\partial \Gamma} \frac{\partial \Gamma}{\partial y^j} \right) + \frac{\partial}{\partial \Gamma} \left( \frac{\partial f}{\partial y^j} + \frac{\partial f}{\partial \Gamma} \frac{\partial \Gamma}{\partial y^j} \right) \frac{\partial \Gamma}{\partial y^i}$$

$$= \frac{\partial^2 f}{\partial y^i \partial y^j} + \frac{\partial^2 f}{\partial y^i \partial \Gamma} \frac{\partial \Gamma}{\partial y^j} + \frac{\partial f}{\partial \Gamma} \frac{\partial^2 \Gamma}{\partial y^i \partial y^j}$$

$$+ \frac{\partial^2 f}{\partial \Gamma \partial y^i} \frac{\partial \Gamma}{\partial y^j} + \frac{\partial^2 f}{\partial \Gamma^2} \frac{\partial \Gamma}{\partial y^i} \frac{\partial \Gamma}{\partial y^j}$$

So, by (1)

$$\alpha_{ij} \frac{\partial^2 f}{\partial y^i \partial y^j} + \beta^i \frac{d}{dy^i} f + ef$$

$$= \left( \alpha_{ij} \frac{\partial \Gamma}{\partial y^i} \frac{\partial^2 f}{\partial \Gamma^2} + 2\alpha^{ij} \frac{\partial \Gamma}{\partial y^j} \frac{\partial^2 f}{\partial y^i \partial \Gamma} + M \frac{\partial f}{\partial \Gamma} + N(f) \right)$$

$$= 4\Gamma \frac{\partial^2 f}{\partial \Gamma^2} + 4y^\sigma \frac{\partial^2 f}{\partial \Gamma \partial y^\sigma} + M \frac{\partial f}{\partial \Gamma} + N(f).$$
Now applying the above formula to $H_k$, we have

$$-A^*_y H_k = \sum_{i=0}^{k} \frac{\Gamma}{4} (t - \tau)^{i-2-m/2} \exp \left( -\frac{\Gamma}{4(t-\tau)} \right) y_i$$

$$+ \sum_{i=0}^{k} (t - \tau)^{i-1-m/2} \exp \left( -\frac{\Gamma}{4(t-\tau)} \right) \left\{ y^\sigma \frac{\partial u_i}{\partial y^\sigma} + \frac{M}{4} u_i - N(u_{i-1}) \right\}$$

$$\quad - (t - \tau)^{k-m/2} \exp \left( -\frac{\Gamma}{4(t-\tau)} \right) N(u_k)$$

where $u_{-1} = 0$, $N(u_{-1}) = 0$. Since

$$-\frac{\partial}{\partial \tau} H_k = \sum_{i=0}^{k} (t - \tau)^{i-1-m/2} \exp \left( -\frac{\Gamma}{4(t-\tau)} \right) u_i \left( -\frac{m}{2} + i + \frac{\Gamma}{4(t-\tau)} \right)$$

the theorem will be proved if we can choose $u_i$ satisfying the relations

$$y^\sigma \frac{\partial u_i}{\partial y^\sigma} + \left( -\frac{m}{2} + i + \frac{M}{4} \right) u_i = N(u_{i-1}),$$

$u_i(P, Q)$ being $C^\infty$ in a vicinity of $P$ with $u_{-1} = 0$ and $u_0(P, P) = 1$. To see that we may choose such $u_i$, put $y^\sigma = \eta^\sigma s$ and transform the equation as an ordinary differential equation in $s$ containing the parameters $\eta$:

$$\frac{du_i}{ds} + \left( -\frac{m}{2} + i + \frac{N}{4} \right) u_i = N(u_{i-1}).$$

Choose

$$u_0 = \exp \left\{ -\int_0^s s^{-1} \left( -\frac{m}{2} + \frac{M}{4} \right) ds \right\}.$$

$u_i$ is $C^\infty$ near $s = 0$, because of the order relation $M = 2m + o(s)$. Define $u_i$ successively by the formula

$$u_i(P, Q) = u_0 s^{-1} \int_0^s s^{-1} u_0^{-1} N(u_{i-1}) ds \quad (i = 1, 2, \ldots, k).$$
Lecture 23

1 Integration of the Fokker-Planck equation
(Contd.) Differentiability and representation
of the operator-theoretical solution

\[ f(t, x) = (T_t f)(x), \quad f \in L_1(R) \]

**Lemma 1.1.** Let \( h(x, \tau) \) be \( C^\infty \) in \((x, \tau) x \in R, t \geq \tau \geq 0\), and vanish outside a compact coordinate neighbourhood of \( P \) (independent of \( \tau \)). Then

\[
\int_R f(y, t) h(y, t) dy = \int_R f(y, o) h(y, o) dy + \int_0^\tau d\tau \int_R \left\{ \partial_\tau f(y, \tau) h(y, \tau) + f(y, \tau) \frac{\partial h(y, \tau)}{\partial \tau} \right\} dy
\]

where \( \partial_\tau f(y, \tau) = \text{strong lim}_{\delta \to 0} \{ f(y, \tau + \delta) - f(y, \tau) \} \delta^{-1} \).

**Proof.** \( f(y, \tau) h(y, \tau) \) is weakly differentiable with respect to \( \tau \) in \( L_1(R) \) and the weak derivative is

\[
\partial_\tau f(y, \tau) h(y, \tau) + f(y, \tau) \frac{\partial h(y, \tau)}{\partial \tau}
\]

\( \square \)
Corollary. We have
\[ \int_R f(y, t) h(y, t) dy = \int_R f(y, o) h(y, o) dy \]
\[ + \int_0^t d\tau \left\{ \int_R f(y, \tau) \left( \frac{\partial h(y, \tau)}{\partial \tau} + A^*_y h(y, \tau) \right) \right\} dy \]

Proof. By Lemma [1.1] the right hand side is
\[ = - \int_0^t d\tau \left\{ f(y, \tau) \left( - \frac{\partial h(y, \tau)}{\partial \tau} - A^*_y h(y, \tau) \right) \right. \]
\[ - h(y, \tau) \left( \partial_, f(y, \tau) - \bar{A}_y f(y, t) \right) \}
\[ = \int_0^t d\tau \int_R \left\{ \partial_, f(y, \tau) h(y, \tau) + f(y, \tau) \frac{\partial h(y, \tau)}{\partial \tau} \right\} dy \]
\[ + \int_0^t d\tau \int_R \left\{ f(y, \tau) A^*_y h(y, \tau) - h(y, \tau) \bar{A}_y f(y, \tau) \right\} dy. \]

\[ \square \]

We have, by the definition of the smallest closed extension \( \bar{A} \),
\[ \int_R \{ f(y, \tau) A^*_y h(y, \tau) - h(y, \tau) \bar{A}_y f(y, \tau) \} dy \]
\[ = \lim_{k \to \infty} \int_R \{ f_k(y, \tau) A^*_y h(y, \tau) - h(y, \tau) A^*_y f_k(y, \tau) \} dy \]
where \( s - \lim f_k = f, s - \lim A^*_y f_k = \bar{A} f \). The integral on the right is zero, by Green’s formula and the fact that \( h \) vanishes near the boundary \( \partial R \).

We take for \( h(y, \tau) \) the function
\[ h(y, \tau) = h(Q, \tau) = H_k(P, Q, t + \epsilon - \tau) \delta(P, Q) \delta(P_o, P); \]

here \( P_o \) is a point of \( R, \epsilon \) a positive constant and \( \delta(P, Q) = \alpha(r(P, Q)) \)
where \( \alpha(r) \) is \( C^\infty \) function of \( r \) such that \( 0 \leq \alpha(r) \leq 1, \alpha(r) = 1 \) for \( r \leq 2^{-1} \eta \) and \( = 0 \) for \( r \geq \eta \). \( \eta > 0 \) is chosen so small that the point \( Q \)
satisfying \( \delta(P_o, P) \delta(P, Q) \) are contained in a compact coordinate neighbourhoud of \( P_o \). We then have

\[
f(Q, t)H_k(P, Q, \varepsilon)\delta(P_o, P)\delta(P, Q)dQ \\
= f(Q, 0)H_k(P, Q, t + \varepsilon)\delta(P_o, P)\delta(P, Q)dQ \\
- \int_0^\varepsilon d\tau \int f(Q, \tau)K_k(P, Q, t + \varepsilon - \tau)dQ
\]  

where

\[
K_k(P, Q, t + \varepsilon - \tau) = -\left( \frac{\partial}{\partial \tau} + A^*_Q \right)H_k(P, t + \varepsilon - \delta) \delta(P_o, P) \delta(P, Q)
\]  

If \( k \) is chosen such that \( k - \frac{m^2}{2} \geq 2 \), then by lemma 1.1, \( K_k(P, Q, t + \varepsilon - \tau) \) is for \( r(P_o, P) \leq 2 - \frac{1}{\eta} \), devoid of singularity even if \( t + \varepsilon - \tau = 0 \). We now show that the left side of (2) tends as \( \varepsilon \downarrow 0 \) to \( f(P, t) \) in the vicinity of \( P_o \).

\[
\int_R \delta(P_o, P)dP|f(Q, t)H_k(P, Q, \varepsilon)\delta(P, Q)dQ \\
- \delta(P, t) \int H_k(P, Q, \varepsilon) \delta(P, Q)dQ|
\]

\[
\leq C \int_{(P, Q) \leq 2\eta} \int_{r(P, Q) \leq \eta} \left| f(Q, t) - f(P, t) \right|H_k(P, Q, \varepsilon)dQ \\
\leq C_1 \int \cdots \int \left( \int |f(z + \varepsilon^{frac{12}{3}} \xi, t) - f(z, t)| d\xi_{\frac{4}{3}} \right) d\xi_1 \cdots d\xi_n
\]

\((z^1 \cdots z^m)\) and \((z^1 + y^1, \ldots, z^m + y^m)\) are coordinates of \( P \) and \( C, C_1 \) are constants. The inner integral on the right converges as \( \varepsilon \downarrow 0 \), to zero boundedly by Lebesgue’s theorem.

There exists therefore a sequence \( \{\varepsilon_i\} \) with \( \varepsilon_i \downarrow 0 \) such that

\[
f(P, t) \lim_{i \to \infty} \int_R H_k(P, Q, \varepsilon_i) \delta(P, Q)dQ = \int_R f(Q, 0)H_k(P, Q, t)\delta(P, Q)dy \\
- \int_0^\varepsilon d\tau \int_R f(Q, \tau)K_k(P, Q, t + \varepsilon - \tau)dQ
\]
almost everywhere with respect to \( P \) in the vicinity of \( P_0 \). Hence \( f(P, t) \) may be considered to be continuously differentiable once in \( t > o \) and twice in \( P \) in vicinity of \( P_0 \) if

\[
\lim_{\varepsilon \downarrow 0} \int_R H_k(P, Q, \varepsilon) \delta(P, Q) dQ
\]

is positive and twice continuously differentiable in \( P \) in the vicinity of \( P_0 \). Now,

\[
\lim_{\varepsilon \downarrow 0} \int H_k(P, Q, \varepsilon) \delta(P, Q) dQ = -\lim_{\varepsilon \downarrow 0} \int e^{m/2} \exp \left( \frac{\Gamma(P, Q)}{4} \right) dQ
\]

for \( \xi > 0 \). Hence, putting

\[
ds^2 = \varphi_{ij}(y)dy^i dy^j, \quad y^i = e^{1/2} \xi^i, \lim_{\varepsilon \downarrow 0} \int R H_k(P, Q, \varepsilon) \delta(P, Q) dQ
\]

\[
= \lim_{\varepsilon \downarrow 0} \int \cdots \int \exp(-\alpha_{ij}(0) \xi^i \xi^j) \varphi(0)^{1/2} d\xi^1 \cdots d\xi^n
\]

\[
= \pi^{m/2} (\varphi(0))^{1/2} (\alpha(0))^{1/2}
\]

\[
= \pi^{m/2} (g(P))^{1/2} / (\alpha(P))^{1/2}
\]

where \( g(P) = \det(G_{ij}(P)) \) and \( \alpha(P) = \det(\alpha_{ij}(P)) \).

Thus in the vicinity of \( P_0, f(P, t) \) is equivalent to

\[
\pi^{m/2} g(P)^{-1/2} \alpha(P)^{1/2} \left\{ \int_R f(Q, 0) H_k(P, Q, t) \delta(P, Q) dQ
\right. \]

\[
- \int_0^t d\tau \int_R f(q, \tau) K_k(P, Q, t-\tau) dQ
\]

So it is differentiable once in \( t \) and twice in \( P \). Moreover, we have \( |f(P, t)| \leq \text{Const} \ |f(Q, 0)| \). Therefore there exists \( \rho(P, Q, t) \) bounded in \( Q \), such that

\[
f(P, t) = \int \rho(P, Q, t) f(Q, 0) dQ.
\]
Bibliography

[1] F.V. Atkinson:

[2] N. Dunford:

[3] W. Feller:
   (iii) Two singular diffusion problems, Ann. of Math. 54 (1951), 173-182.
BIBLIOGRAPHY

(vii) Semi-groups of transformations in general weak topologies, Ann. of Math. 57 (1953), 287-308.


[4] K. O. Friedrichs:


[5] M. Fukamiya:

(i) On one-parameter groups of operators, Proc. Imp. Acad. Tokyo, 16 (1940), 262-265.

[6] I. Gelfand:


[7] E. Hille:


(viii) Le probleme abstrait de Cauchy, Rendiconti del Seminario Matematico di Univ. di Torino, 12 (1952-1953), 95-103.


[8] K. Ito:


[9] S. Ito:


[10] T. Kato:


[12] P.D. Lax:

[13] P. D. Lax and A. N. Milgram:

[14] S. Minakshisundaram:

[15] S. Minakshisundaram and A. Pleijel:

[16] L. Nirenberg:

[17] F. Perrin:
   (i) Etude mathematique du mouvement brownien de rotation, Ann. Ec. norm. sup. 45 (1928), 1-51.

[18] R. S. Phillips:

[19] G. de Rham:


[20] F. Riesz and B. Sz. Nagy:

(i) Lecons d’analyse fonctionnelle, Budapest, 1952.

[21] M.H. Stone:

(i) Linear transformations in Hilbert space III, Proc. Nat. Acad. Sci. 16(1930), 172-175.

[22] L. Schwartz:


[23] K. Yosida:


(iv) Integration of Fokker-Planck’s equation in a compact Riemannian space, Arkiv for Matematik, 1 (1949), 244-253.


