

Ph129 Lecture Notes  
Part B -

B. Vectors, Matrices and Functional Analysis. Application to Integral Transforms and Differential and Integral Equations

B.1 Motivation

- (i) So far we have been doing ~~from~~ rather grimy - albeit at times useful - mathematics. Now we will approach the same and other questions from a more general ~~over~~ point of view.

References:

Introduction → M. and W. chapter 6 (must read this - we will go further)  
 Lucid, old fashioned language and long → Courant and Hilbert, "Methods of Mathematical Physics - vol 1" - chapters 1 and 2.  
 → D. and K. - chapters 2 and 3.

not perfect but best.

We will briefly analyse a very familiar and simple construct and show how many seeming very complicated problems can be handled

in the same language.

- (ii) Consider plain old 3-space which people used to live afore the advent of special relativity. A point in our space is denoted by a vector  $\underline{x} = (x_1, x_2, x_3)$ . The following concepts - are among those that - have proved useful.

(a) Linear structure: if  $\underline{x}$  and  $\underline{y}$  are vectors, so is  $\alpha \underline{x} + \beta \underline{y}$ :  $\alpha, \beta$  (real) live numbers.

(b) 3 axes! Any vector can be written as a l.c. of 3 independent vectors.

$$\underline{x} = x_1 \underline{e}_1 + x_2 \underline{e}_2 + x_3 \underline{e}_3.$$

$$\underline{e}_1 = (1, 0, 0)$$

$$\underline{e}_2 = (0, 1, 0)$$

$$\underline{e}_3 = (0, 0, 1).$$

(3)

(c) One can change bases ~~bases~~ by an (orthogonal or otherwise: orthogonal if metric preserved) transformation to any 3 linearly independent vectors  $\underline{e}'_1$ ,  $\underline{e}'_2$  and  $\underline{e}'_3$ . So the same vector

$$\underline{x} = x'_1 \underline{e}'_1 + x'_2 \underline{e}'_2 + x'_3 \underline{e}'_3$$

(d) This change of co-ordinate systems can be represented by a  $3 \times 3$  matrix  $A$ .  
if  $|e'_i\rangle = B|e_j\rangle$  defines operator  $B$

$$\langle e'_j | e'_i \rangle = \langle e'_j | B | e_i \rangle = B_{ji}$$

$$\text{or } |e'_i\rangle = \sum_j B_{ji} |e_j\rangle$$

$$\text{i.e. } \sum_{j=1}^3 x_j |e_j\rangle = \sum_{j,k=1}^3 x'_k B_{kj} |e_j\rangle$$

$$\text{or } x_j = \sum_i B_{ji} x'_i$$

$$\text{or } x'_i = A_{ij} x_j \quad (A = B^{-1})$$

$A$  is a matrix - or more abstractly -

a linear operator. Considering  $\underline{x}$  and  $\underline{x}'$  as different vectors referred to same base

$$\underline{x}' = A \underline{x}$$

and  $A$  is a linear operator that takes a vector in one <sup>linear</sup> space

(4)

into a vector in <sup>another</sup> (generally different) linear vector space. In the special case of square matrices, the initial and final spaces are identical.

(e) Linear Equations:

The equations

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

can be summarized by the matrix equation

$$A\underline{x} = \underline{b}.$$

Given  $\underline{b}$  and  $A$  we want find a vector  $\underline{x}$ , such that linear operator  $A$  acting on  $\underline{x}$  gives  $\underline{b}$ . Solution is formally

$$\underline{x} = A^{-1}\underline{b}$$

but it only exists under certain conditions. In this case  $\det(A) \neq 0$  is sufficient (not necessary of course).

(f) metric / Norm / Scalar Product:

Our 3-space has a distance defined in it: namely

$$\underline{x}^2 = x_1^2 + x_2^2 + x_3^2.$$

In fact we have the more general concept of a scalar product

$$\underline{x} \cdot \underline{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

Lapsing into Dirac notation, we often write:

$$\underline{x}^2 = \langle x|x \rangle.$$

$$\underline{x} \cdot \underline{y} = \langle x|y \rangle.$$

(iii) Now let's show that the problems we face in studying differential and integral equations are in fact very similar.

Thus let  $C[a, b]$  be the set of all continuous functions  $f(x)$  of a real variable  $x$  in the closed range  $[a, b]$ . Then

(a)  $C[a, b]$  is a linear vector space, where a "vector" is a function  $f(x)$ . So if we have two vectors  $f \sim f(x)$  and  $g \sim g(x)$  then clearly  $\alpha f(x) + \beta g(x)$  ( $\alpha, \beta$  real numbers) is also a continuous function over  $[a, b]$  and hence  $\int$  a well defined meaning for  $\alpha f + \beta g$  as a vector in  $C[a, b]$ .

(b) clearly  $C[a, b]$  is a bit different from 3-space. certainly we can't expand all its members in terms of a finite number of members of  $C[a, b]$ . i.e.

$$f(x) \neq \sum_{n=1}^N f_n e_n(x)$$

for some  $N$ ,  $e_n(x)$  and all  $f(x)$ .

However similar expansions are familiar.

i.e. Legendre expansion

$$f(x) = \sum_{n=0}^{\infty} a_n P_n(x) \quad \begin{matrix} [a, b] \\ = [-1, 1] \end{matrix}$$

or Fourier series

$$f(x) = \sum_{n=0}^{\infty} (b_n \cos nx + c_n \sin nx)$$

are well-known for some ranges  $[a, b] = [-\pi, \pi]$  and "suitable"  $f$ .

$[a, b] \times \underline{\text{so}}$  as upper limit on  $n$  is  $\infty$ ,

$C[a, b]$  is an infinite dimensional vector space.

So the "Integral Transform" problem:

$$f(x) = \sum_{n=0}^{\infty} f_n e_n(x) \quad \text{--- (B1/2)}$$

$e_n(x)$  = Legendre, Trigonometric, Bessel functions etc., is analogous to the choice of basis in 3-space.  $e_n(x)$  are complete if (B1/2) possible for all  $f \in C[a, b]$ . In 3-space we know that any 3 independent vectors are complete. Basic problem for integral transforms is find out for a given vector space  $C[a, b]$ , whether or not set is complete.

- (c) After choice of basis we studied operators. what is the analogue of this in  $C[a, b]$ ?

Any differential equation can be represented by an operator. Thus

(9)

take  $y'' + y = q(x)$   
 and write it as  $Ly = q$

where  $y \in C[a, b]$  is vector  $y(x)$   
 $q \in \quad \quad \quad q(x)$

and operator  $L$  takes  $y(x)$  to  
 new vector  $y''(x) + y(x)$ .

If differential equation is linear,  
 then so is its associated operator  $L$ .

Now  $Ly = q$  is clearly  
 analogous to  $Ax = b$  in 3-space.  
 So solution of a D.E. boils down to  
 finding operator  $L^{-1}$  such that

$y = L^{-1} q$  just as we  
 solved  $Ax = b$  as  $x = A^{-1} b$

Now differential operators are  
 tricky as need conditions on  $y$   
 for (a)  $L$  to be defined even i.e.

$y''$  must exist and (p) even if  $q = Ly$  defined,  $q$  may be outside original space. e.g.  $y(x) = \theta(x)x^2$  has discontinuous  $y''(x)$ .

For this reason, study of D.E. is more difficult than that of integral equations

e.g.

$$f(x) = u(x) + \int_a^b k(x,t) f(t) dt$$

is in operator form:

$$f = u + K f$$

where  $K$  is a linear operator.

For reasonable  $k(x,t)$ ,  $K f \in C[a,b]$  if  $f \in C[a,b]$ .

(c) Finally we can define a norm and even a scalar product.

$$\langle f | g \rangle = \int f(x) g(x) dx \quad \sim (B1.2)$$

This is crucial because convergence criterion e.g.

$$f(x) = \sum_{n=0}^{\infty} a_n P_n(x) \quad \sim (B1.3)$$

really means

$$\left| f - \sum_{n=0}^N a_n P_n \right| < \epsilon \quad \forall N > N(\epsilon) \quad \sim (B1.4)$$

i.e. partial sums converge to  $f$ .  
without this (B1.3) doesn't mean anything.

We can take various choices for  $\| \cdot \|$

in (B1.4) e.g. (B1.2)

$\|f(x)\|^2 = \int_a^b dx f^2(x)$  gives  
"convergence in mean", while

$$\|f(x)\| = \max_{a \leq x \leq b} |f(x)|$$

gives more familiar uniform convergence.

So functional analysis is study of infinite dimensional vector spaces and involves bases, operators and norms. we will start by defining linear spaces abstractly, then revise properties of finite dimensional spaces, then treat bases in  $\infty$ -dimensional spaces (w) and apply to integral transforms; then we treat operators in  $\infty$ -dimensional spaces and apply to differential and integral equations.

## B2. Basic Definitions: Vector Spaces, Metric Spaces, Banach Spaces and Hilbert Spaces

### (i) Vector Spaces:

A vector space is a set of objects  $|a\rangle, |b\rangle$  satisfying:  
A: operation of addition of objects is defined and multiplication by scalars. (latter are in practice real or complex numbers). Thus

- (a) If  $|a\rangle, |b\rangle \in S$ , then  $(|a\rangle + |b\rangle) \in S$ .
- (b) If  $|a\rangle \in S$  and  $\alpha$  a scalar  $\alpha|a\rangle \in S$ .
- (c)  $\exists$  an element  $|0\rangle \in S$  such that for any  $|a\rangle \in S$  one has  $|a\rangle + |0\rangle = |a\rangle$ .
- (d) For any  $|a\rangle \in S$ ,  $\exists$  vector  $|-a\rangle \in S$  such that  $|a\rangle + |-a\rangle = |0\rangle$

B: The addition and multiplication defined above, satisfy the following properties:

- (a)  $|a\rangle + |b\rangle = |b\rangle + |a\rangle$  : addition of vectors is commutative.
- (b)  $(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle + |c\rangle)$  : addition of vectors is associative.
- (c)  $1 \cdot |a\rangle = |a\rangle$
- (d)  $\alpha \cdot (\beta |a\rangle) = (\alpha\beta) |a\rangle$  : associative law of multiplication by scalars.
- (e)  $(\alpha + \beta) |a\rangle = \alpha |a\rangle + \beta |a\rangle$  : distributive law of multiplication by sum of scalars.
- (f)  $\alpha (|a\rangle + |b\rangle) = (\alpha |a\rangle) + (\alpha |b\rangle)$  : distributive law for vectors.

These 10 axioms define a linear or vector space. They seem trivial; many trivial results can be proved from them!

(i)  $|0\rangle$  is unique:

Proof: let  $|a\rangle + |0\rangle = |a\rangle$  : all  $a$

and  $|a\rangle + |0'\rangle = |a\rangle$  : some  $b$

add.  $1-b \rangle$  to both sides of last

$$1-b \rangle + (1-b \rangle + 10' \rangle) = (1-b \rangle + 1-b \rangle)$$

$$\text{or } 10 \rangle + 10' \rangle = 10 \rangle$$

Take  $1a \rangle = 10' \rangle$  in First equation

$$10' \rangle + 10 \rangle = 10' \rangle$$

commutativity  $\Rightarrow 10 \rangle = 10' \rangle$  i.e  $10 \rangle$  unique

(ii)  $1-a \rangle = (-2) 1a \rangle$

put  $(1 + (-2)) 1a \rangle = 10' \rangle$

$$\begin{aligned} \text{Then } 10' \rangle + 1a \rangle &= [1 + (-2) + 2] 1a \rangle \\ &= 1a \rangle \end{aligned}$$

$$\Rightarrow 10' \rangle = 10 \rangle \quad \text{by (i)}$$

So  $10 \rangle = 1a \rangle + (-2) 1a \rangle$

$$\Rightarrow 1-a \rangle = (-2) 1a \rangle \quad \text{Q.E.D.}$$

(reader can prove  $1-a \rangle$  unique, see problems).

## Examples of Vector Spaces

- (a) Euclidean Space  $E^{(n)}$  : set of ~~any~~  $n$ -tuples of  $n$  real numbers  $\underline{x} = (x_1, x_2, \dots, x_n)$  with obvious multiplication by scalar
- $$\alpha(x_1, x_2, \dots, x_n) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$$
- and addition of vectors

$$\underline{x} + \underline{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

This is a vector space with the scalars in definition being taken from (field of) real numbers. One can also use complex  $n$ -tuples to get complex vector space  $C_{\mathbb{C}}^{(n)}$  with scalars from (field of) complex numbers.

- (b) our old space  $C[a, b]$  of continuous functions over range  $x \in [a, b]$

$$f \rightarrow f(x), \quad \alpha f \rightarrow \alpha f(x)$$

$$f + g \rightarrow f(x) + g(x) \text{ of course.}$$

We have already abstracted the notion of linearity. Now let us abstract the notion of distance.

### (ii) Metric Spaces

A set  $R$  is called a metric space if a real (positive) number  $p(a, b)$  is associated with any pair of its elements  $a, b \in R$  and if

$$(a) \quad p(a, b) \geq 0$$

$$(b) \quad p(a, b) = p(b, a)$$

$$(c) \quad p(a, b) = 0 \text{ only when } a \equiv b$$

(d)  $p(a, b) + p(b, c) \geq p(a, c)$ . This is triangle inequality - obvious for distances in a plane:



$p$  is called the metric on  $R$ .

(6)

$\rho$  is abstraction of distance in Euclidean Space  $E^n$

$$\rho(x, y) = [(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2]^{1/2}$$

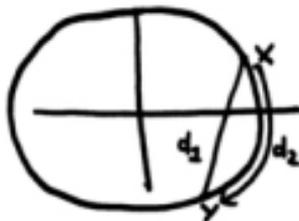
$E^n$  is also a vector space: we can find a metric space which isn't a vector <sup>space</sup> by taking some subset in  $E^n$  e.g. in  $E^{(2)}$



Also note that the same set can have different metrics defined on it.

(a) let  $R$  be set in  $E^{(2)}$  with

$$x^2 + y^2 = 1.$$



$d_1$  is usual Euclidean distance

$d_2$  is arc-length along circle (by shorter of two routes!). Easily check  $d_2$  satisfies axioms.

(b) Back to our faithful  $C[a, b]$ .

$$g_1(f, g) = \max_{a \leq x \leq b} |f(x) - g(x)|$$

$$g_2(f, g) = \int_a^b dx |f(x) - g(x)|^2$$

In metric spaces one can define concept of convergence. Thus  $x_n \rightarrow x$  in  $R$  if  $\rho(x, x_n) \rightarrow 0$  as  $n \rightarrow \infty$ .

One also defines a Cauchy sequence  $\{x_n\}$  in  $R$  by the condition that if  $\epsilon$  given  $\epsilon > 0$ ,  $\exists N$  s.t.  $\rho(x_n, x_m) < \epsilon$  for  $n, m > N$ . Finally we have

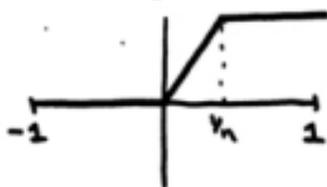
crucial definition: a metric space  $R$  is complete if every Cauchy

Sequence has a limit.

$E^{(n)}$  is complete : arrows geometrically

For instance set of numbers  $x$  in  $0 < x < 1$  is an incomplete metric space. Sequence  $(1, 1/2, 1/3, 1/4 \dots)$  is Cauchy but has no limit. In this case we can make the space complete by adding zero element. This is always true : any metric space can be embedded in a larger metric space in which it is complete. We will later try to realise this completion for  $C[a, b]$  with metric  $d_p$  for it is clearly not complete. Thus take  $[a, b] = [-1, 1]$ . Define :

$$f_n(x) = \begin{cases} 1 & \frac{1}{n} \leq x \leq 1 \\ nx & 0 \leq x \leq \frac{1}{n} \\ -1 & -1 \leq x \leq 0 \end{cases}$$



$$\begin{aligned} \text{Then } \int_{-1}^{+1} |f_n(x) - f_m(x)|^2 & \quad (\text{norm wolog}) \\ &= \int_0^{1/n} |f_n(x) - f_m(x)|^2 \\ &\leq \frac{1}{n} \rightarrow 0 \text{ as } n, m \rightarrow \infty. \end{aligned}$$

So  $f_n$  is a Cauchy sequence in  $C[-1, +1]$  with norm  $g_2$  but it converges to

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & -1 < x < 0 \end{cases}$$

which is discontinuous and so outside space. Note that  $C$  is complete.

wrt norm  $\rho_2$  (in this case  $f_n$  is no longer Cauchy and so above counter example is invalid).

### (iii) Banach ~~metric~~ Spaces

Above we were discussing metric spaces: many of which were also vector spaces. These are blessed with separate names.

Definition: A normed vector space/space has all the axioms of a vector space plus

(a) To every  $|x\rangle \in N$  is associated a real number  $\|x\|$  which satisfies ...

(b)  $\|\alpha x\| = |\alpha| \|x\|$

(c)  $\|x\| = 0$  iff  $|x\rangle = |0\rangle$

(d)  $\|x+y\| \leq \|x\| + \|y\|$  : this is triangle inequality in a new guise.

replacing  $\rho(x, y)$  by  $\|x - y\|$   
 we see that a normed vector space  
 is a metric space. We define the  
 important concept of completeness for  
 a normed vector space and:

Definition: A complete normed vector  
 space is called a Banach Space.

#### (iv) Hilbert Spaces

Definition: A complex vector space  $V$   
 is called an inner product space if there  
 is a complex valued function (  
 written  $\langle x | y \rangle$ ) of vectors  $|x\rangle$  and  $|y\rangle$   
 which satisfies

$$(a) \quad \langle x | x \rangle \geq 0$$

$$(b) \quad \langle x | x \rangle = 0 \quad \text{iff} \quad |x\rangle = |0\rangle$$

$$(c) \quad \langle x | y + z \rangle = \langle x | y \rangle + \langle x | z \rangle$$

$$(d) \quad \langle x | \alpha y \rangle = \alpha \langle x | y \rangle$$

$$(e) \quad \langle x | y \rangle = \overline{\langle y | x \rangle}$$

This notation and properties should be familiar from quantum mechanics. (inner product called scalar product by physicists).  
Theorem:  $\|x\| = \langle x|x \rangle^{1/2}$  is a norm.

Proof: The only nontrivial condition to be satisfied is the  $\Delta$  inequality. First we prove the Cauchy-Schwarz inequality which states  $\langle x|x \rangle \langle y|y \rangle \geq |\langle x|y \rangle|^2$  (B2.1) and follows at once from:

$$\langle x + \alpha y | x + \alpha y \rangle \geq 0 \quad \text{all } \alpha$$

$$\text{i.e. } \langle x|x \rangle + \alpha^* \langle y|x \rangle + \alpha \langle x|y \rangle + |\alpha|^2 \langle y|y \rangle \geq 0$$

$$\text{Put } \alpha = \langle y|x \rangle \beta \quad \beta \text{ real}$$

$$\langle x|x \rangle + 2\beta |\langle y|x \rangle|^2 + \beta^2 \langle y|y \rangle |\langle y|x \rangle|^2 \geq 0$$

which is quadratic in  $\beta$  and must have no real roots. i.e. " $b^2 \leq 4ac$ "

$$\text{or } 4 |\langle y|x \rangle|^4 \leq 4 \langle y|y \rangle |\langle y|x \rangle|^2$$

which implies (B2.3). Given this

we immediately calculate:

$$\begin{aligned}
 \|x+y\|^2 &= \langle x|x \rangle + 2 \operatorname{Re} \langle x|y \rangle + \langle y|y \rangle \\
 &\leq \langle x|x \rangle + 2 |\langle x|y \rangle| + \langle y|y \rangle \\
 &\leq \langle x|x \rangle + 2 \langle x|x \rangle^{1/2} \langle y|y \rangle^{1/2} + \langle y|y \rangle \\
 &= (\langle x|x \rangle^{1/2} + \langle y|y \rangle^{1/2})^2 \\
 &= (\|x\| + \|y\|)^2 \qquad \text{Q.E.D.}
 \end{aligned}$$

So an inner product space is a normed vector space. It is sometimes called a prehilbert space. However a complete inner product space is called a Hilbert space. Note that any Hilbert space is a Banach space but not vice versa:

e.g.

(a)  $\left\{ \begin{matrix} \mathbb{R}^{(n)} \\ \mathbb{C}^{(n)} \end{matrix} \right\}$  is a Hilbert space with

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n. \quad (82.2)$$

Note any finite dimensional space is always complete.

(b) Remember a given set can have many different metrics. e.g. consider the space  $\mathbb{E}^{(n)}$  consisting of the same vectors

$(x_1, \dots, x_n)$  but with norm:

$$\|x\| = \max_{1 \leq i \leq n} |x_i| \quad (B2.2)$$

to replace  $\|x\|^2 = x_1^2 + x_2^2 + \dots + x_n^2$  (B2.3)

associated with  $\mathbb{E}^{(n)}$ . (B2.4) can be extended

to give the scalar product (B2.2) : (B2.3)

has ~~not~~ no such natural extension. In fact

one can show there isn't one: however

$\mathbb{E}^{(n)}$  is complete and so is not a Hilbert

Space but rather a Banach Space let us

prove this for the slightly more

interesting case  $n = \infty$ , when we get

Space  $l_\infty$ . This is set of all sequences

$(x_1, \dots, x_n, \dots)$  of real numbers satisfying

$$\sum_i |x_i| < \infty. \quad (B2.5)$$

l.u.b. stands for least upper bound and is a technically accurate version of "max" for infinite sequences. Consult your favorite analysis book if you're uncertain on this point. (B2.5) indicates that it is sensible ~~to~~ to define a norm by

$$\|x\| = \text{l.u.b. } |x_i| \quad (\text{B2.6})$$

(B2.6) clearly makes  $\ell_\infty$  a normed vector space: take, for instance, the  $\Delta$  inequality

$$\|x+y\| = \text{l.u.b. } |x_i + y_i|$$

$$\text{but for each } i, |x_i + y_i| \leq |x_i| + |y_i|$$

$$\text{so } \|x+y\| \leq \text{l.u.b. } [|x_i| + |y_i|]$$

$$\text{clearly } \leq \text{l.u.b. } |x_i| + \text{l.u.b. } |y_i|$$

$$= \|x\| + \|y\|$$

Q.E.D.

So we need only show  $\ell_\infty$  is complete. To do this suppose  $x_{i,n}$  is a

Cauchy sequence in  $l_\infty$ .

$$x_n = (x_{1n}, \dots, x_{in}, \dots)$$

Then for any  $\epsilon > 0$ ,  $\exists N$

$$\|x_n - x_m\| < \epsilon \quad n, m > N$$

In particular this implies

$$|x_{in} - x_{im}| < \epsilon \quad n, m > N \quad \text{each } i \text{ (82.7)}$$

So for each  $i$ ,  $x_{in}$  is a Cauchy sequence in  $\mathbb{R}$  (real numbers). So as  $\mathbb{R}$  complete,  $\exists y_i$

$$x_{in} \rightarrow y_i \quad \text{as } n \rightarrow \infty$$

$$\text{let } y = (y_1, y_2, \dots, y_i, \dots)$$

we must show that a)  $y$  is in  $l_\infty$  and

b)  $\|x_n - y\| \rightarrow 0$  as  $n \rightarrow \infty$ . Letting  $m \rightarrow \infty$

in (82.7) we find

$$|x_{in} - y_i| < \epsilon \quad n > N$$

take any such  $N$

$$\cancel{\text{but}} |y_i| \leq |x_{in}| + |x_{in} - y_i|$$

$$\leq \epsilon + \|x_n\| < \infty$$

i.e.  $y$  is in  $l_\infty$ .

Again.

$$\|y - x_n\| = \lim_i |x_{in} - y_i| < \epsilon \quad n > N$$

i.e.  $\|y - x_n\| \rightarrow 0$  as  $n \rightarrow \infty$

i.e.  $l_\infty$  is complete Q.E.D.

So  $l_\infty$  is indeed a Banach

Space

- (c) we could have defined a more natural extension of  $\mathbb{R}^{(n)}$ ,  $\mathbb{C}^{(n)}$  i.e. the set of all sequences  $\{x_1, x_2, \dots\}$  with  $\sum_{i=1}^{\infty} |x_i|^2 < \infty$

This is called  $l_2$  and with

Scalar product

$$\langle y | x \rangle = \sum_{i=1}^{\infty} y_i^* x_i \tag{82.8}$$

It can be shown to be a Hilbert Space (problem set)

It fact we shall see that all infinite dimensional (separable) Hilbert Spaces are isomorphic to  $l_2$ . So it is only Hilbert space!

(d) ~~(a)~~ <sup>Similarly</sup> One can show that  $C[a, b]$  with metric  $\rho_2$  is not a Hilbert Space. [one of problem set] -

(e) ~~(b)~~ But  $C[a, b]$  with metric  $\rho_2$  can be extended to an inner product space i.e.

$$\langle f | g \rangle = \int_a^b f^*(x) g(x) dx \quad (B2.9)$$

is a scalar product. we have yet to get a Hilbert Space as Space not complete.

(v) Completion of  $C[a, b]$

The fact that  $C$  is not complete causes great damage in the succeeding sections. Courant and Hilbert overcome this by restricting the set  $C[a, b]$  to ~~the~~ <sup>any</sup> subset  $E(a, b)$  of equicontinuous functions. A set of continuous functions  $f(x)$  in  $[a, b]$  is equicontinuous if given  $\epsilon > 0$ ,  $\exists \delta(\epsilon)$  that is independent of  $f$  such that

$$|f(x_1) - f(x_2)| < \delta \quad \text{for all } f \quad (B2.10)$$

with  $|x_1 - x_2| < \epsilon$  and  $x_1, x_2 \in [a, b]$ .

This gives convergence properties for functions similar to those for real numbers. Unfortunately the condition clearly violates the linear structure of our space.

Thus (B2.10) clearly cannot be satisfied for all functions gotten by multiplication of one of them by arbitrary constants.

So the alternative approach - avoided by C. and H. because of the necessity of Lebesgue integration - is to extend the set C. This can either be done formally:

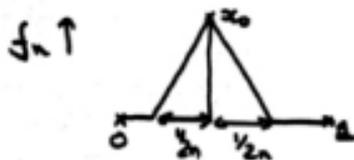
i.e. use general procedure for completing metric space. This involves taking space C and taking any Cauchy sequence  $\{f_j\}$ . If it has a limit in C fine. Otherwise we invent a limit for it and consider the new space consisting of old points plus new "ideal" element. The process is continued

An analogous argument defines real numbers from rationals.

until all limits found. Some of the new limits will be "equal" (i.e.  $g(x,y) = 0 \Rightarrow |x| = |y|$ ): in this case we must consider them equal - or more precisely take equivalence classes wrt. metric.

However this is not very transparent.

Better is to explicitly extend the class of functions considered; from our previous example ~~we~~ (page 19) we certainly need discontinuous functions. In particular we <sup>have</sup> limit of



which is 0 at everywhere in  $[a, b]$  (here  $[0, 2]$ ) except  $x_0$  where it is 1. Doing this denumerably often. (possible

with limit denumerable  $n \rightarrow \infty$ ) we can get Dirichlet's function in our class of admissible functions. This is defined in  $[0, 1]$  and

$$\chi(x) = \begin{cases} 0 & x \text{ is irrational.} \\ 1 & x \text{ is rational.} \end{cases}$$

(remember rationals are denumerable).

Given all these horrible functions we must be able to integrate them so as to be able to form scalar products as in (82.9). Fortunately the answer is classical - the completion of  $C[a, b]$  is the set of all Lebesgue integrable functions in  $[a, b]$ . If you know what Lebesgue integration is - fine. Otherwise one need only

- (a) browse through D. and K. pages 184-89
- (b) Note for sensible functions,

Lebesgue integral = usual (Riemann) integral. We can thus denote it by the same thing.

(c) for continuous functions

$$\int_a^b f^2(x) = 0 \text{ implies } f(x) \equiv 0.$$

This is not true of all Lebesgue integrable functions.

e.g.  $\int_a^b \chi(x) = 0$  even though

$\chi(x)$  not equal to zero at denumerably many points. This is important because

one axiom of a Hilbert space is that  $\|x\| = 0 \iff |x\rangle = |0\rangle$ . It follows

that we must regard  $f_1(x)$  and  $f_2(x)$  as equal if  $\int_a^b (f_1(x) - f_2(x))^2 dx = 0$ .

We say  $f_1 = f_2$  almost everywhere

or  $f_1 = f_2$  a.e. Alternately  $f_1 \neq f_2$   
on a set of measure zero.

To summarize, and generalize  
we have

### Riesz-Fischer Theorem

Let  $L_w^2 [a, b]$  be set of Lebesgue  
integrable functions with weight  $w(x)$   
over  $[a, b]$ . ( $w(x) > 0$ )

i.e.  $\int_a^b w(x) |f(x)|^2 dx$  exists  $< \infty$ .

Define equivalence classes on this  
set by  $f_1 \approx f_2$  if  $f_1 = f_2$  a.e.

Then the set of equivalence classes  
is a Hilbert Space with Scalar product

$$\langle f | g \rangle = \int_a^b w(x) f^*(x) g(x) dx \quad (2.12)$$

As  $L_w^2 [a, b]$  is a Hilbert Space  
it is complete. When we say

a sequence  $f_k(x) \rightarrow f(x)$  in this space,  
we say "convergence is in mean":

$$\text{i.e. } \lim_{k \rightarrow \infty} \int_a^b w(x) (f_k(x) - f(x))^2 dx = 0 \quad (B2.12)$$

This is not the same as

$$\lim_{k \rightarrow \infty} f_k(x) = f(x) \quad [\text{this}$$

corresponds to convergence w.r.t norm  
 $\lim_{a \rightarrow x \rightarrow b} |f(x)|$  in our Banach space].

In particular (B2.12) is still true if  
we change  $f(x)$  on any set of measure  
zero e.g. add  $137\chi(x)$ .

### 83 General Theory of orthonormal sets in Hilbert Spaces

(i) In a finite dimensional Hilbert space of dimension  $n$ , any vector can be expressed in terms of a set  $|g_i\rangle$

$$|x\rangle = \sum_{i=1}^n \alpha_i |g_i\rangle \quad (83.1)$$

where the  $|g_i\rangle$  are linearly independent i.e.  $\nexists$  no linear rel<sup>n</sup>  $\sum_{i=1}^n \beta_i |g_i\rangle = 0$  unless all the  $\beta_i = 0$ . We can make this basis more convenient by the Schmidt orthogonalization method.

$$\text{This puts } |e_2\rangle = |g_2\rangle / \sqrt{\langle g_2 | g_2 \rangle}$$

$$|e_2\rangle = \frac{1}{N_2} [ |g_2\rangle - \langle e_1 | g_2 \rangle |e_1\rangle ]$$

where  $N_2$  is chosen so that  $\langle e_2 | e_2 \rangle = 1$   
or inductively

~~$$|e_2\rangle = \frac{1}{N_2} [ |g_2\rangle - \langle e_1 | g_2 \rangle |e_1\rangle ]$$~~

$$|e_l\rangle = \frac{1}{N_l} \left[ |g_l\rangle - \sum_{i=1}^{l-1} \langle e_i | g_l \rangle |e_i\rangle \right] \quad (B3.2)$$

( $N_l \neq 0$  from linear independence of  $|g_l\rangle$ )  
clearly from (B3.2) for  $l < l$

$$\begin{aligned} \langle e_j | e_l \rangle &= \frac{1}{N_l} \left[ \langle e_j | g_l \rangle - \sum_{i=1}^{l-1} \langle e_i | g_l \rangle \delta_{ij} \right] \\ &= 0. \quad (\text{by inductive hypothesis.}) \end{aligned}$$

This is the Schmidt method by  
constructing from  $|g_i\rangle$  a new set  
 $|e_j\rangle$  that is orthonormal

$$\text{i.e. } \langle e_i | e_j \rangle = \delta_{ij}$$

and so that the  $|g_i\rangle$  can be expressed  
linearly in terms of  $|e_i\rangle$  so that (B3.1)  
by be rewritten

$$|x\rangle = \sum_{i=1}^n \alpha_i |e_i\rangle \quad \text{--- (B3.3)}$$

dotting both sides with  $\langle e_j |$  gives

$$\alpha_i = \langle e_j | x \rangle \quad \text{--- (B3.4)}$$

➡ Note, in passing, that

$$\begin{aligned}\langle x|x \rangle &= \sum_{i=1}^n x_i^* \sum_{j=1}^n x_j \langle e_i | e_j \rangle \\ &= \sum_{i=1}^n |x_i|^2\end{aligned}$$

$$\text{or } \langle x|x \rangle = \sum_{i=1}^n |\langle e_i | x \rangle|^2 \quad \text{--- (83.5)}$$

which is called Parseval's relation

(ii) Now let's try to generalize this to an  $\infty$  dimensional space.

Take any vector  $|g_1\rangle$  in  $H$ : if  $H$  is not 1 dimensional,  $\exists$  a vector  $|g_2\rangle$  in  $H$  linearly independent of  $|g_1\rangle$ . The process may be continued to find a denumerable independent set  $|g_1\rangle \dots |g_n\rangle \dots$

The Schmidt process may be used to orthogonalize these and find an infinite set  $|e_i\rangle$  such that

$$\langle e_i | e_j \rangle = \delta_{ij}.$$

Remembering (1), take any vector  $|x\rangle$  in  $H$  and put

$$|x_R\rangle = \sum_{i=1}^R \langle e_i | x \rangle |e_i\rangle = \sum_{i=1}^R x_i |e_i\rangle \quad (B3.6)$$

Now we proved that

$$|\langle x | x_R \rangle|^2 \leq \langle x | x \rangle \langle x_R | x_R \rangle \quad (B3.7)$$

Plugging (B3.6) into (B3.7) gives

$$\left( \sum_{i=1}^R |x_i|^2 \right)^2 \leq \langle x | x \rangle \left( \sum_{i=1}^R |x_i|^2 \right)$$

$$\text{or } \sum_{i=1}^R |x_i|^2 \leq \langle x | x \rangle$$

Letting  $R \rightarrow \infty$

$$\sum_{i=1}^{\infty} |x_i|^2 \leq \langle x | x \rangle \quad (B3.8)$$

which is Bessel's inequality which among other things proves that series  $\sum_{i=1}^{\infty} |x_i|^2$  converges. Comparing with (B3.5) we see we get in finite dimensional case: equality in (B3.8) if we include all members of ~~the~~ a basis set; The inequality if we

leave out some members.

Definition : an orthonormal set  $|e_i\rangle$  is

said to be closed if

$$\sum_{i=1}^{\infty} |\langle e_i | x \rangle|^2 = \langle x | x \rangle \quad \text{for all } |x\rangle \in H.$$

Theorem:  $|e_i\rangle$  is closed iff

$$|x\rangle = \sum_{i=1}^{\infty} \langle e_i | x \rangle |e_i\rangle \quad \text{for all } |x\rangle \quad (83.9)$$

Proof: (83.9) means that if

$$|x_R\rangle = \sum_{i=1}^R \langle e_i | x \rangle |e_i\rangle, \quad \text{then } |x_R\rangle \rightarrow$$

$|x\rangle$  as  $R \rightarrow \infty$ . Suppose  $|e_i\rangle$  closed, then

$$\begin{aligned} \|x_R - x\|^2 &= \langle x_R - x | x_R - x \rangle \\ &= \langle x_R | x_R \rangle - 2 \operatorname{Re} \langle x | x_R \rangle + \langle x | x \rangle \\ &= \sum_{i=1}^{\infty} |\langle e_i | x \rangle|^2 - \sum_{i=1}^R |\langle e_i | x \rangle|^2 \rightarrow 0 \quad \text{as } R \rightarrow \infty \end{aligned}$$

The converse is similar.

Definition : an orthonormal set  $|e_i\rangle$

is said to be complete if the only vector orthogonal to all its members is the null vector.

Theorem :  $\{e_i\}$  is complete iff

$$|x\rangle = \sum_{i=1}^{\infty} \langle e_i | x \rangle |e_i\rangle \text{ for all } x \quad (B3.9)$$

Proof : define  $|x_k\rangle$  as usual, and take  $k > l$

$$\|x_k - x_l\|^2 = \sum_{n=k+1}^l |\langle x_n | e_n \rangle|^2 \quad (B3.20)$$

But from Bessel's inequality  $\sum_{n=1}^{\infty} |\langle x | e_n \rangle|^2$

is a convergent series. So  $|\langle x | e_n \rangle|^2$  is a

Cauchy sequence and rhs of (B3.20) is  $< \epsilon$

for  $k, l > N(\epsilon)$ . Thus  $|x_k\rangle$  is a Cauchy

sequence in  $H$  and as  $H$  is complete, it has

a limit vector  $|y\rangle$  in  $H$ . We must now

show  $|y\rangle = |x\rangle$ . It is easy to see that

$$\langle e_i | y \rangle = \lim_{k \rightarrow \infty} \langle e_i | x_k \rangle \quad (\text{problem sheet})$$

But for  $k > i$ ,  $\langle e_i | x_k \rangle = \langle e_i | x \rangle$

So  $\langle e_i | y \rangle = \langle e_i | x \rangle$  all  $i$

i.e.  $|x\rangle - |y\rangle$  is orthogonal to all  $|e_i\rangle$

i.e.  $|x\rangle = |y\rangle$  by completeness of  $|e_i\rangle$

The converse is trivial

So the concepts of completeness and closure for orthonormal bases are equivalent in Hilbert Spaces and are again equivalent to the existence of expansions:

$$|x\rangle = \sum_{i=1}^{\infty} \langle e_i | x \rangle |e_i\rangle \quad \text{for all } x$$

Note  $\langle x | x \rangle = \sum_{i=1}^{\infty} |\langle e_i | x \rangle|^2$  (B3.12)

$$\langle y | x \rangle = \sum_{i=1}^{\infty} \langle y | e_i \rangle \langle e_i | x \rangle$$

(ii) The one and only Hilbert Space

Definition: A metric space is separable if it has a countable dense subset.

(A subset  $S$  is dense in  $H$ , if given  $|x\rangle$  in  $H$  and  $\epsilon > 0$ ,  $\exists |y\rangle$  in  $S$  with  $\|x - y\| < \epsilon$ ). For instance, the real numbers are separable because the rationals are a countable dense subset.

Given a dense subset in  $H$ , we can throw out non linearly independent

members and orthogonalize the remainder with the Gram-Schmidt procedure. This gives an orthonormal set  $|e_i\rangle$  such that the sums  $\sum_{i=1}^n \alpha_i |e_i\rangle$  for all  $n$  are dense in  $H$ . Clearly  $|e_i\rangle$  are a complete set and (B3.12) holds.

Theorem : Any separable Hilbert Space  $H$  is isomorphic either to  $\mathbb{C}^n$  or  $l_2$ .

Proof : The case when  $H$  is finite dimensional and isomorphic to  $\mathbb{C}^n$  is obvious. Thus take an infinite dimensional separable Hilbert Space : from the above argument, there is a countable complete orthonormal basis  $|e_i\rangle$ . Set an isomorphism  $H$  to  $l_2$  by :  $|x\rangle \in H$  into  $(\alpha_1, \alpha_2 \dots \alpha_n \dots) \in l_2$

where  $\alpha_i = \langle e_i | x \rangle$   
 n.b.  $|e_i\rangle = (0, 0, 0, \dots, 1, 0, 0, 0, \dots)$  is complete in  $l_2$ .

Parseval's relation ensures that  $\sum_{i=1}^{\infty} |a_i|^2 < \infty$   
and that we have a true member of  $\ell_2$ .

The only thing to be checked is that  
given  $(\beta_1 \dots \beta_n \dots) \in \ell_2$  with  $\sum_{i=1}^{\infty} |\beta_i|^2 < \infty$   
 $\exists$  a corresponding  $|x\rangle = \sum_{i=1}^{\infty} \beta_i |e_i\rangle$ .

This follows as in theorem on page 42.

Put  $|x_n\rangle = \sum_{i=1}^n \beta_i |e_i\rangle$ .  $|x_n\rangle$  is (from convergence of  $\sum_{i=1}^{\infty} |\beta_i|^2$ ) a Cauchy sequence and hence has a limit in  $H$  - called  $|x\rangle$ .

g. e. d.

Most Hilbert spaces are separable - certainly  $L^2_w [a, b]$  is - and we now turn to a proof of this.

### B4 Weierstrass's Theorem

In this section, we show how to construct a complete orthonormal set  $\{e_n\}$  in  $L^2_w[a, b]$ . We shall assume that  $w(x)$  is a nice positive function and  $\int_a^b w(x) x^n dx$  exists for all  $n$  (e.g. this rules out  $w(x) \propto (x-a)^{\alpha}$  :  $\alpha < -2$ ).

We assumed the Riesz-Fischer theorem that  $L^2_w[a, b]$  was the completion of  $C[a, b]$ . Being the completion, it is clear from the formal method of completing a metric space that  $C[a, b]$  is dense in  $L_w[a, b]$ . It follows that if we can find a basis for  $C[a, b]$ , it will also be one for  $L_w[a, b]$ . So in order to prove that ~~the~~  $\{e_n\} = x^n$  is a complete basis for  $L^2_w[a, b]$ , it is

Sufficient to prove:

Weierstrass's Approximation Theorem:

Let the function  $f(x)$  be continuous on the finite closed interval  $[a, b]$ . For any  $\epsilon > 0$ ,  $\exists$  a positive integer  $n$  and a corresponding polynomial  $p_n(x)$  of the  $n$ 'th degree such that

$$|f(x) - p_n(x)| < \epsilon \quad \text{for any } x \in [a, b] \quad (B4.1)$$

we give two proofs of this: both use explicit construction:

(i) In the first method, replace  $[a, b]$  by  $[0, 1]$  w.o.l.o.g., and introduce the Bernstein polynomials.

$$B_n(f, x) = \sum_{k=0}^n \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$$

These can be shown to satisfy (B4.1) (with large enough  $n$  and  $p_n(x) = B_n(x)$ ). This can be seen qualitatively from the result:

$$\binom{n}{R} x^R (1-x)^{n-R} \approx \exp \left[ -\frac{(R-k_0)^2}{2\sigma^2} \right] + O(1/n) \quad (84.2)$$

where  $k_0 = xn$  and  $\sigma^2 = x(1-x)n$ .

Only terms with  $R = k_0 \pm c\sqrt{n}$  are important in sum: these ~~terms~~ give  $f(x \pm c/\sqrt{n}) \rightarrow f(x)$ . (84.2) is usual result of ~~the~~ central limit theorem in statistics. Although this is an intuitively pleasing derivation (for those who know statistics), it appears that a rigorous derivation is not easy with central limit theorem. Rather one must replace latter by rigorous bounds.

(ii) D. and K. (pages 199 → 202) and C. and H. (pages 65 → 68) use the same method which depends on the following lemma.

Lemma: For any  $1 > \delta > 0$  and with

$$A_n(\delta) = \int_{\delta}^1 (1-y^2)^n dy \quad (84.3)$$

one has  $\lim_{n \rightarrow \infty} \left\{ A_n(\delta) / A_n(0) \right\} = 0$

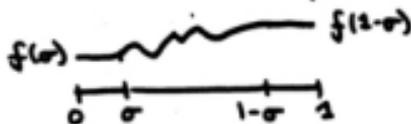
Proof:

$$A_n(\delta) = \int_{\delta}^1 (1-y^2)^n < \int_{\delta}^1 (1-\delta^2)^n \leq (1-\delta^2)^n$$

$$A_n(0) \geq \int_0^1 (1-y)^n = \frac{1}{n+1} \Rightarrow \text{result.}$$

Proof of w.A.T.

First map  $[a, b]$  into  $[0, 1-\sigma]$  :  $\sigma > 0$



and extend continuous functions over  $[0, 1-\sigma]$  to be continuous

over  $[0, 1]$  by making them constant in  $[0, \sigma]$  and  $[1-\sigma, 1]$ .

According to a standard (but non-obvious) theorem,  $\exists \delta(\epsilon)$

$$|f(x+y) - f(x)| < \epsilon/2 \quad \text{all } x \text{ in } [0, 1] \\ \text{for } |y| < \delta \quad (B4.4)$$

Define

$$p_n(x) = \frac{1}{2A_n(\delta)} \int_0^1 f(z) [1 - (x-z)^2]^n dz$$

This is a polynomial of degree  $2n$ .

Note  $[1 - (x-z)^2]^n$  peaks at  $x=z$  and dies away rapidly thereafter. Write:

$$\begin{aligned} \frac{1}{2A_n(\delta)} \int_0^1 f(z) [1 - (x-z)^2]^n dz \\ = \int_{-x}^{-\delta} \frac{f(x+y)}{2A_n(\delta)} [1 - y^2]^n dy + \int_{-\delta}^{+\delta} \text{ditto} + \int_{\delta}^{1-x} \text{ditto} \\ = I_1 + I_2 + I_3 \end{aligned}$$

$$\text{Now } I_2 = \frac{f(x)}{2A_n(\delta)} \int_{-\delta}^{\delta} (1 - y^2)^n dy + I_4$$

$$|I_4| < \frac{\epsilon}{2A_n(\delta)} \int_{-\delta}^{\delta} (1 - y^2)^n dy \cdot \epsilon/2$$

$$\text{or } I_2 = f(x) \left\{ 1 - \frac{A_n(\delta)}{A_n(0)} \right\}$$

$$|I_2| < \epsilon/2$$

meanwhile, down on the form,

$$\begin{aligned} I_1 + I_3 &< \frac{\max |f|}{2A_n(0)} \left\{ \int_{-1}^{-\delta} (1-y^2)^n dy + \int_{\delta}^1 (1-y^2)^n dy \right\} \\ &= \max |f| \frac{A_n(\delta)}{A_n(0)} \end{aligned}$$

$$\begin{aligned} \text{or } |I_1 + I_2 + I_3 - f(x)| &< \epsilon/2 + 2 \max |f| \frac{A_n(\delta)}{A_n(0)} \\ &< \epsilon \text{ for large } n \end{aligned}$$

$$\text{as } \frac{A_n(\delta)}{A_n(0)}$$

Q.E.D.

This completes proof that  $1, x, x^2, \dots$  are complete in  $L^2_w[a, b]$ . It only works if  $[a, b]$  finite: C. and H. (pages 95-97). Show that polynomials are complete in  $L^2_{w=\exp(-x)}[0, \infty]$

and  $L^2_{w=\exp(-x^2)}[-\infty, \infty]$ . Their proof is not complete however and one should really read Szegő, "Orthogonal Polynomials", pages 104-105.

Now we have shown that  $1, x, \dots, x^n, \dots$  is a complete set in  $L_w[a, b]$  for any reasonable  $w(x)$ . This shows that all these Hilbert spaces (for different  $w(x)$   $a$  and  $b$ ) are the same and isomorphic to  $l_2$ . However there is still a lot to be gained by considering them separately: things may be obvious in one space and not other. (co-ordinates changes are often vital in simple finite dimensional problems). In particular if we use Gram-Schmidt orthogonalization with the different scalar products, we will get different linear orthonormal combinations  $|e_n\rangle$ .

All will be polynomials satisfying

$$\int_a^b w(x) p_n(x) p_m(x) dx = \delta_{nm}.$$

we now turn to a consideration of the various  $w(x)$  and their associated polynomials. In each case expansion theorems of (B3) are valid and we need not repeat them - they are of course important: our unified treatment has allowed us to dispense with individual formulation. For instance  $w(x) = 1$  gives us Legendre polynomials with  $[a, b] = [-1, 1]$  - use Schmidt orthogonalization to show this!

Then any  $f(x) = \sum c_n \sqrt{\frac{2n+1}{2}}$   $P_n(x)$

$$c_n = \int_{-1}^{+1} f(x) P_n(x) \sqrt{\frac{2n+1}{2}} dx$$

is "partial wave analysis"

## 85 General Properties of Orthogonal Polynomials

Read Chapter III of Szegő: "Orthogonal Polynomials"; it is quite amazing how much one can prove from the simple condition

$$\int_a^b w(x) p_n(x) p_m(x) dx = 0 \quad n \neq m \quad (85.1)$$

we will not specify overall normalization - as this usually done by other criteria - and so we avoid  $\delta_{nm}$  in (85.1).

### (i) Recurrence Relation (D. and K. pages 208-209).

We are all familiar with:

$$(n+2) P_{n+2}(x) = (2n+1)x P_{n+1}(x) - n P_n(x) \quad (85.2)$$

and I pointed out this was a special case of a recurrence relation valid for all hypergeometric functions. The latter in fact was quite powerful as it held for complex  $n$ . However we now go the other way around our mathematical circle and show that all orthogonal polynomials obey such a recurrence relation. This is different - for

instance we could find no hint of orthogonality relas for general hypergeometric equations. As a continuing aside, we note that properties of classical special functions can be regarded as amalgam of properties true for all hypergeometric (confluent possibly) functions (valid all n) and those true for orthogonal polynomials (integer n).

Theorem:  $\{p_n(x)\}$  satisfy

$$p_{n+2}(x) - (A_n x + B_n) p_n(x) + C_n p_{n-1}(x) = 0 \quad (B.5.3)$$

where  $k_n =$  coefficient of  $x^n$  in  $p_n(x)$

$k'_n =$  " "  $x^{n-1}$  in  $p_n(x)$

$$h_n = \int_a^b p_n^2(x) w(x) dx$$

then  $A_n = k_{n+1}/k_n$

$$B_n = k_{n+1}/k_n \left[ k'_{n+1}/k_{n+1} - k'_n/k_n \right]$$

$$C_n = h_n/h_{n-2} \cdot k_{n+1} k_{n-1}/k_n^2$$

The statement of the pudding is longer than the proof....

Proof

$p_n(x)$  is orthogonal to  $p_0(x) \dots p_{n-1}(x)$   
 " " " "  $x p_0(x) \dots x p_{n-2}(x)$

i.e.  $x p_n(x)$  " " "  $p_0(x) \dots p_{n-2}(x)$

Then  $p_{n+1}(x) - \frac{k_{n+1}}{k_n} x p_n(x)$

is orthogonal to  $p_0(x) \dots p_{n-2}(x)$  and has no coefficient of  $x^{n+1}$ . It is thus a linear combination of  $p_n(x)$  and  $p_{n-1}(x)$ . This proves (B5.3). We need just specify  $B_n$  and  $C_n$ .

(a) compare coefficients of  $x^n$

$$k'_{n+1} - A_n k'_n = B_n k_n$$

i.e.  $B_n = k_{n+1}/k_n \left( k'_{n+1}/k'_{n+1} - k'_n/k_n \right)$  as requested

(b)  $\int w(x) x p_{n-1}(x) p_n(x) dx$  on (B5.3). This gives

$$0 - A_n \int w(x) x p_{n-1}(x) p_n(x) dx + C_n h_{n-1} = 0$$

Now  $x_{p_{n-1}}(x) = \frac{R_{n-1}}{R_n} p_n(x) + \text{Series in lower } p_m\text{'s}$

$$So - A_n R_{n-1}/R_n h_n + C_n h_{n-2} = 0$$

$$i.e. C_n = h_n/h_{n-2} R_{n+1} R_{n-1}/R_n^2 \quad \text{p.e.d.}$$

### (ii) Christoffel - Darboux Formulae

Let  $x, y$  be any points. Define Kernel  $K$  by:

$$K_n(x, y) = \sum_{k=0}^n P_k(x) P_k(y) / h_k \quad \text{--- (B5.4)}$$

If  $\pi_n(x)$  is any polynomial of degree  $n$ .

$$\int_a^b K_n(x, y) \pi_n(x) w(x) dx = \pi_n(y) \quad \text{(B5.5)}$$

### Theorem

$$K_n(x, y) = \frac{R_n}{R_{n+1} R_{n-1} h_n} \left[ \frac{P_n(y) P_{n+1}(x) - P_n(x) P_{n+1}(y)}{x-y} \right] \quad \text{(B5.6)}$$

This is easily proved by induction by substituting (B5.3) for  $P_{n+1}(x)$ ,  $P_{n+1}(y)$  on right hand side of (B5.6).

Theorem

Consider any polynomial  $\pi_n(x)$  of degree  $n$  and any complex point  $z_0$ . Then the  $\pi_n(x)$  that maximizes

$$|\pi_n(z_0)|^2$$

subject to  $\int_a^b |\pi_n(x)|^2 w(x) dx = 1$  (B5.7)

is given by:  $\pi_n(x) = e^{i\theta} (K_n(z_0^*, z_0))^{-1/2} K_n(z_0^*, x)$  for any  $\theta$  real  $\theta$ .

Proof: expand

$$\pi_n(x) = \sum_{m=2}^n \lambda_m p_m(x) / \sqrt{h_m}$$

Then constraint is  $\sum_{m=2}^n |\lambda_m|^2 = 1$ .

But by Cauchy's inequality

$$|\pi_n(z_0)|^2 \leq \sum_{m=2}^n |\lambda_m|^2 \sum_{m=2}^n |p_m(z_0)|^2 / h_m$$

$$= K_n(z_0^*, z_0)$$

But this bound is attained for (B5.7) <sub>g.e.d</sub>

### (iii) Zeros of Orthogonal Polynomials

Theorem: The zeroes of the orthogonal polynomials  $p_n(x)$  are real, distinct and all lie in the interior of  $[a, b]$ .

Proof: Let  $x_0$  be any zero of  $p_n(x)$

Let  $q(x) = p_n(x)/(x-x_0)$  a polynomial of degree  $n-1$  (as  $p_n$  has real coefficients its zeroes are complex conjugate).

$$\text{Then } \int_a^b w(x) p_n(x) q(x) dx = 0$$

$$\text{or } \int_a^b w(x) (x-x_0) |q(x)|^2 dx = 0$$

$$\text{or } x_0 = \frac{\int_a^b w(x) x |q(x)|^2 dx}{\int_a^b w(x) |q(x)|^2 dx}$$

$$\Rightarrow a < x_0 < b.$$

If  $p_n(x)$  had a double zero at  $x_0$ ,  $p_n(x)/(x-x_0)^2 = r(x)$  is a polynomial of degree  $n-2$ .

Then  $\int w(x) p_n(x) r(x) dx = 0$

or  $\int w(x) \left\{ \frac{p_n(x)}{(x-x_0)} \right\}^2 dx = 0$

a contradiction Q.E.D.

Theorem: The zeroes of  $p_n(x)$  and  $p_{n+1}(x)$  interlace. (use recurrence relation (B5.3) to prove)

Theorem: Between any two zeroes of  $p_n(x)$ ,  $\exists$  a zero of  $p_{n+1}(x)$ ,  $n > 0$ . (Szegő, page 45)

Theorem: Let  $(\alpha, \beta)$  be any subinterval of  $[a, b]$ . For some  $n$ ,  $p_n(x)$  vanishes in  $(\alpha, \beta)$ . (see problem set).

(iv) Approximation Theory

The first is well known for all orthonormal systems. It is part of problem set.

Theorem let  $d = \| f(x) - \pi_n(x) \|$  for fixed  $f(x)$  and any polynomial  $\pi_n(x)$  of degree  $n$ . [remember  $d^2 = \int_a^b \omega(x) dx (f(x) - \pi_n(x))^2$ ]

Then  $d$  is minimized by the choice

$$\pi_n(x) = \sum_{R=0}^n \alpha_R P_R(x) \quad (B5.8)$$

$$\alpha_R = \frac{1}{\sqrt{h_R}} (P_R, f)$$

i.e. usual "convergence in mean" formula. This states that "convergence in mean"

$$f(x) = \sum_{R=0}^{\infty} \alpha_R P_R(x) \text{ is the best in}$$

if cutoff at any order  $R=n$ , the best you can do with polynomials of this order.

Theorem: Let  $f(x)$  be continuous and  $\pi_n(x)$  determined by (B5.8). Then  $f(x) - \pi_n(x)$  changes sign at least  $n+2$  times in  $[a, b]$  or else vanishes identically.

Proof:  $q(x) = f(x) - \pi_n(x)$  is orthogonal to  $p_0(x) \dots p_n(x)$ .

$$\int_a^b q(x) w(x) dx = 0$$

ensures one zero. Label zeroes  $x_1, x_2$

~~standing from~~  $\therefore x_l$  in  $[a, b]$  ordered in increasing size.

$$\int q(x) (x-x_1) \dots (x-x_l) w(x) dx = 0$$

gives another zero if  $l \leq n$ .  $\therefore l \geq n+2$ .

Q.E.D.

Now we come to a very important result for numerical integration. One use of polynomial approximation to functions is that one can manipulate former easily - in particular one can often integrate them when one can't integrate function.

A typical numerical integration formula is

$$\int_a^b w(x) f(x) dx = \sum_{i=1}^n p_i f(x_i) \quad (B5.9)$$

where  $p_i$  are independent of  $f$ . If say  $x_i$  are prescribed equi-spaced points,

$$\begin{aligned} x_1 &= a \\ x_2 &= a+h \\ &\vdots \\ x_n &= a+(n-1)h=b \end{aligned}$$

Then we can choose the  $n$  parameters  $p_i$  so that (B5.9) is exact for  $f(x)$  polynomials of degree  $n-1$  or else. Equivalently we ~~find~~ find the unique polynomial  $\pi_{n-1}(x)$  of degree  $n-1$  that takes on the  $n$ -values  $f(x_i)$  at  $x=x_i$ . These give so-called Newton-Cotes formulae of which the most famous is Simpson's rule

$$\int_a^b f(x) dx = \frac{(b-a)}{6} \left[ f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]$$

which is  $n=3$  in above. [it happens also to be exact for cubics].

Gaussian integration formulae are gotten by allowing  $x_i$  to vary in (85.9). This gives us  $n$  more parameters and we can get a better integration formula (requiring fewer evaluations of  $f$  i.e. fewer dollars), which is valid for polynomials of degree  $2n-1$ .

Theorem : If  $x_1 \dots x_n$  are the zeroes of  $p_n(x)$ ,  $\exists$  real positive numbers  $\lambda_i$  such that 
$$\int_a^b \pi_{2n-1}(x) w(x) dx = \sum_{i=1}^n \lambda_i \pi_{2n-1}(x_i) \quad (85.10)$$

and, in fact,  $\lambda_i = 1/K_n(x_i, x_i) \quad (85.11)$

Proof: Taking  $\pi_{2n-1}(x) = p_n(x) \begin{matrix} 1 \\ x \\ \vdots \\ x^{n-1} \end{matrix} \Rightarrow$   $n$  linear equations for  $\lambda_i$  for  $i=1, \dots, n$   
 $\Rightarrow \lambda_i p_n(x_i) = 0$

for which integral is zero, we easily see that  $x_i$  must be zeroes of  $p_n(x)$ .

To find  $\lambda_i$ , let

$$p(x) = (x-x_1) \dots (x-x_{i-1})(x-x_{i+1}) \dots (x-x_n)$$

the product of all the zeroes of  $p_n(x)$

Save  $x_i$ . Then  $K_n(x_i, x)$  ~~is~~  $g(x)$  is a polynomial of degree  $2n-1$  for which (B5.10) must be valid. But on right hand side we only get term with  $x_i$  contributing i.e.

$$\lambda_i g(x_i) K(x_i, x_i) = \int w(x) K(x_i, x) g(x) dx$$

but from (B5.5), we find integral is just  $g(x_i)$ .

So  $\lambda_i^{-1} = K(x_i, x_i)$  as desired  $\square$ .

we'll come back to this when we do numerical analysis. Turns out - although Gaussian integration is much prettier - Simpson's rule is usually preferable as it allows iterative use on the computer. Still important to note that theory of orthogonal polynomials has many important practical applications - the latter often use  $w(x) \neq 1$ . This is not a mathematical necessity - proper ~~or~~ choice of  $w(x)$  in a

given problem can make or break the assumption that  $f(x)$  is well approximated by a low order polynomial.

Another minimizing problem is to find the polynomial of degree  $n$  that minimizes

$$\max_{a \leq x \leq b} \{ |f(x) - \pi_n(x)| w(x) \}$$

i.e. back to our old norm in Banach Space  $C[a, b]$  of continuous functions. This is harder mathematical problem and in the case  $w(x) = 1$  leads to Tchebichef polynomials. We will do this in numerical analysis section of course. [the hardness reflects fact that Banach spaces have fewer properties than Hilbert spaces i.e. they lack a scalar product].

### 86 Particular cases of Orthogonal Polynomials

(a) In the last two sections we learnt what marvellous things (expansion, gaussian integration) one could do with orthogonal polynomials: however we didnt actually show that any existed! In principle given any weight function  $w(x)$  we can use the Schmidt orthogonalization method to

construct  $p_n(x)$  inductively by:

$$p_0(x) = \text{const (e.g. 1)} \tag{86.1}$$

$$\int_a^b w(x) p_n(x) p_m(x) dx = 0, n \neq m$$

(see the beginning of section 83 for details). Although this is what one does for the (in practice mythical) general case, ~~these~~ the usual orthogonal polynomials can be discussed more elegantly.

In particular we follow D. and K.

(page 203 onwards) and ~~also~~ treat the so-called classical orthogonal polynomials.

(b) Consider the Generalized Rodriguez

Formula: to motivate this remember

the familiar Rodriguez formula

$$P_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} (1 \cdot (1-x^2)^n) \quad (B6.2)$$

applicable to the Legendre polynomials

which are orthogonal over  $[-1, +1]$  with weight function  $w(x) \equiv 1$ . (i.e.  $\int_{-1}^{+1} P_n(x) P_m(x) dx = 2\delta_{nm} / (2n+1)$ ).

Generalize, by writing

$$C_n(x) = \frac{1}{w(x)} \frac{d^n}{dx^n} [w(x) S^n(x)] \quad (B6.3)$$

which is, upto an irrelevant constant,

(B6.2) with  $w=1$  and  $S(x) = 1-x^2$ . We

must impose conditions on (B6.3) to ensure that  $C_n(x)$  are indeed polynomials

and that they have desired orthogonality.

The desired conditions turn out to be:

- (i)  $C_n(x)$  is a first degree polynomial in  $x$ .
- (ii)  $S(x)$  is a polynomial of degree  $\leq 2$  in  $x$  with real roots.
- (iii)  $w(x)$  is real, positive and integrable and satisfies the boundary conditions

$$w(a)S(a) = w(b)S(b) = 0 \quad (B6.4)$$

We postpone the problem of finding functions  $w(x)$  and  $S(x)$  to satisfy all the conditions above. Rather we show that, when we do this, the resultant functions  $C_n(x)$  solve our problem (B6.2) i.e.  $\{C_n(x)\}$  are indeed polynomials of degree  $n$  and moreover they are orthogonal wrt weight function  $w(x)$  over the interval  $[a, b]$ . The proof of both

these follows: first note that putting  $n=1$  in (B6.3) we get:

$$S \frac{dw}{dx} = w \left[ C_1(x) - \frac{dS}{dx} \right]$$

Then if we let  $P(s; k)(x)$  be any polynomial of degree  $\leq k$ , then

$$\begin{aligned} & \frac{d}{dx} (w s^n P(s; k)) \\ &= \frac{dw}{dx} s^n P(s; k) + n w s^{n-1} \frac{ds}{dx} P(s; k) + w s^n \frac{dP(s; k)}{dx} \\ &= w s^{n-1} \left\{ P(s; k) \left[ \underbrace{\frac{s \frac{dw}{dx}}{w} + \frac{ds}{dx} + (n-1) \frac{ds}{dx}}_{= C(s)} \right] + \underbrace{\frac{s \frac{dP(s; k)}{dx}}{dx}}_{\substack{= \text{polynomial degree } \leq 1 \\ = \text{polynomial degree } \leq 2}} \right\} \end{aligned}$$

$$= w s^{n-1} P(s; k+1)$$

or inductively:  $\frac{d^l}{dx^l} (w s^n P(s; k)) = w s^{n-l} P(s; k+l)$   
 where  $l \geq 0$  is any integer. (B6.5)

Now suppose  $m < n$  and let

$$I = \int_a^b P_m(x) C_n(x) w(x) dx = \int_a^b \frac{P_m(x) w \frac{d^n}{dx^n} (w s^n) dx}{w dx^n}$$

Integrate by parts once

$$I = \left[ R_m(x) \frac{d^{n-1}}{dx^{n-1}} (w s^n) \right]_a^b - \int_a^b \frac{dP_m}{dx} \frac{d^{n-1}}{dx^{n-1}} (w s^n) dx \quad (\text{B6.6})$$

But take  $k=0, l=n-1$  in (B6.5)

$$\frac{d^{n-1}}{dx^{n-1}} (wS^n) = w(x)S(x) P_{(n-1)}(x)$$

and so using condition (iii)  $w(a)S(a) = w(b)S(b)$ ,

we see  $\int_a^b$  in (B6.6) is zero. We can continue integration by parts picking up zero surface term, each time and so

$$I = (-1)^n \int_a^b \left[ \frac{d^n}{dx^n} P(x) \right] wS^n dx = 0$$

which shows that  $C_n$  is orthogonal to all polynomials of degree  $< n$ . Again taking  $k=0, l=n$  in (B6.5) we see

$C_n$  is  $P_{(n)}$  and it must in fact be  $P_{(n)}(x)$  [i.e. coefficient of  $x^n$  nonzero] because otherwise it would be a linear combination of  $1, x, \dots, x^{n-1}$  which is impossible as it is orthogonal to all of them. (It is not clear to me that  $C_n(x)$  couldn't be identically zero but in solving (i) .. (iii), we will find this

is not so).

(c) Possible solutions of (B6.3) and conditions (i) to (iii).

Now  $c_{(1)}(x)$  is some linear function  $\alpha x + \beta$  and so we are free to change variables so that  $\alpha x + \beta \rightarrow x$  and so (B6.3) for  $n=1$  becomes a differential equation for  $w(x)$  in terms of  $s(x)$ .

$$\frac{1}{w} \frac{dw}{dx} = - [x + ds/dx] / s \quad \text{--- (B6.7)}$$

$$w(a)s(a) = w(b)s(b) = 0$$

According to condition (ii), we must take in succession  $s =$  zeroth, first and second order polynomials. This is done in detail by D. and K. Here just look at one example

$s = \text{const}(\alpha)$  when (B6.7) becomes

$$\frac{dw}{dx} = -w \frac{x}{\alpha}, \text{ a simple D.E.}$$

with solution  $w(x) = \text{const.} \exp[-x^2/2\alpha]$  (B6.8)

$s = \alpha$  and (B6.8) only satisfy  $s(x)w(x) = 0$

for  $x = \pm\infty$ . So, in this case, we must

take  $a = -\infty$  and  $b = +\infty$ . we can also take  $\alpha = 1/2$ , wolog, by change in the scale of  $x$ . (yet again - we already did it once to get (B6.7)). Then the associated polynomials are, upto a constant, the Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2})$$

To reproduce table 2 in D. and K. page 25,

we summarize the three possible choices for  $S(x)$  as:

Interval $(a, b)$	weight function	$S(x)$	Name
$(-\infty, +\infty)$	$e^{-x^2}$	1	Hermite, $H_n(x)$
$[0, +\infty)$	$x^\nu e^{-x} (\nu > -1)$	$x$	Laguerre, $L_n^\nu(x)$
$[-1, +1]$	$(1-x)^\nu (1+x)^\mu$ ( $\nu, \mu > -1$ )	$(1-x^2)$	Jacobi, $P_n^{(\nu, \mu)}(x)$

Notice that polynomials always have some random conventional normalization so that one should write instead of (B6.3)

$$C_n(x) = \frac{1}{K_n} \frac{d^n}{dx^n} [w(x) S^n(x)] \quad (\text{B6.9})$$

for some constant  $K_n$ .

Note, the Jacobi polynomials have many special cases which are well-known e.g.  $\mu = \nu = 0$  gives us weight function  $w(x) = 1$  and we're back to the Legendre polynomials. Before we delve into specific examples, note that Abramowitz and Stegun p. 773 onwards, chapter 22, has a really comprehensive survey and D. and K. themselves have a water down version which I'm handing out as a student's guide.

Apart from trivial argument changes, the important special cases of Jacobi polynomials are.

Interval	Weight Function	$S(x)$	Name
$[-1, +1]$	$(1-x^2)^{\lambda-1/2}$ , $\lambda > 1/2$	$(1-x^2)$	Gegenbauer, $C_n^\lambda(x)$
"	$\frac{1}{(1-x^2)^{-1/2}}$	"	Legendre, $P_n(x)$
"	$(1-x^2)^{-1/2}$	"	Tchebichef of 1st kind: $T_n(x)$
"	$(1-x^2)^{1/2}$	$(1-x^2)$	Tchebichef of 2nd kind: $U_n(x)$

(d) We proved in (B.5) many general properties of orthogonal polynomials following simply from the orthogonality condition (B.2). The "classical" orthogonal polynomials enjoy several other attributes. In particular differential equations, integral representations and generating functions. Some forms of the latter (especially integral representations) follow from the differential equations (e.g. Jacobi polynomials are hypergeometric functions and we discussed general integral representations for these earlier in course).

### (e) Differential Equations

These are neat because the form can be derived in general. Follow the (usual) leader  $D$  and  $K$ .

Take  $k=n-1$ ,  $l=1$ ,  $n=1$  in (B6.5)

$$\frac{1}{w(x)} \frac{d}{dx} \left[ S w \frac{dC_n}{dx} \right] = P(n) \\ = \sum_{m=1}^n P_m C_m \text{ by completeness}$$

Now multiply both sides by  $w(x) C_k(x)$  and integrate from  $a$  to  $b$ . ( $k \leq n-1$ )

$$\int_a^b C_k(x) \frac{d}{dx} \left[ S w \frac{dC_n}{dx} \right] dx = P_k \cdot h_k$$

$$= \int_a^b \underbrace{\frac{1}{w} \frac{d}{dx} \left[ \frac{dC_k(x)}{dx} S w \right]}_{\text{polynomial of degree } \leq k} C_n(x) dx \text{ by two integrations by parts.}$$

$\leq k$  (take  $l=1$ ,  $n=1$ ,  $k=k-1$  in (B6.5)).

= 0 as  $C_n(x)$  is orthogonal to any polynomial of degree  $\leq n-1$ . ( $k \leq n-1$ ).

So  $P_k = 0$  for  $k \leq n-1$ , and we write

$$\frac{d}{dx} \left[ S w \frac{dC_n}{dx} \right] = -w \lambda_n C_n \quad (\text{B6.10})$$

The constant  $\lambda_n$  has to be determined:

this is detailed in D. and K.

$$\lambda_n = -n \left[ \kappa_1 \frac{dc_1}{dx} + \frac{1}{2} (n-2) \frac{d^2s}{dx^2} \right] \quad (86.11)$$

(f) Generating Functions

These are convenient but have no fundamental importance. A full list is given on page 783 of Abramowitz and Stegun. Generally

$$g(x, z) = \sum_{n=0}^{\infty} a_n C_n(x) z^n \quad (86.12)$$

gives a generating function  $g(x, z)$  for each  $a_n$ : choose  $a_n$  to make  $g$  simple!

An easy example is Tchebichef's first good idea  $T_n(x)$  which is conveniently defined by:

$$T_n(\cos \theta) = \cos n\theta \quad (86.13)$$

Taking  $a_n = 1$

$$\begin{aligned} g(x, z) &= \sum_{n=0}^{\infty} \cos n\theta z^n \\ &= \frac{1}{2} \sum_{n=0}^{\infty} (e^{in\theta} z^n + e^{-in\theta} z^n) \\ &= \frac{1}{2} \left[ \frac{1}{1 - ze^{i\theta}} + \frac{1}{1 - ze^{-i\theta}} \right] \\ &= \frac{1}{2} \left( \frac{2 - 2z \cos \theta}{1 - 2z \cos \theta + z^2} \right) = \frac{1 - xz}{1 - 2xz + z^2} \quad (86.14) \end{aligned}$$

The quantity  $R^2 = 1 - 2xz + z^2$  appears in the generating functions of all the Jacobi polynomials - of which Tchebichef (86.14) is a special case. [This can probably be understood in terms of group-theoretic origin e.g. for integer  $\mu, \nu$  the Jacobi polynomials  $P_n^{(\mu, \nu)}(x)$  are representation functions of the rotation group  $O(3)$ . For instance [M. and W. p. 170]

$$R^{-2} = [1 - 2xz + z^2]^{-1/2} = \sum_{n=0}^{\infty} P_n(x) z^n \quad (86.15)$$

### (h) Integral Representations

These are generally

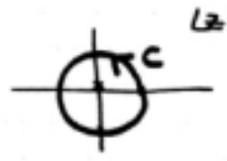
$$C_n(x) = \frac{g_0(x)}{2\pi i} \oint_C [g_2(z, x)]^n g_2(z, x) dz \quad (86.16)$$

where  $C$  is some closed contour in the  $z$  plane.

Given a generating function  $g(x, z)$ , (B6.12),

and suppose  $g$  regarded as a function of  $z$ , it is regular in a neighbourhood of the origin. Consider:

$$\frac{1}{2\pi i} \oint_C z^{-n-1} g(x, z) dz$$



where  $C$  is such that only singularity of integrand is  $z^{-n-2}$ . Then, by usual residue theorem, integral is just an  $T_n(x)$  and we have achieved an integral representation of the type (B6.16). For instance for the Tchebichef polynomials,  $g(x, z)$  satisfies the analyticity conditions if we choose  $C$  so that the zeroes of  $1-2xz+z^2$  lie outside  $C$ . Then

$$T_n(x) = \frac{1}{2\pi i} \oint_C z^{-n} \frac{1}{z} \frac{(1-xz)}{(1-2xz+z^2)} dz \quad (B6.17)$$

Orthogonal Polynomials

Abramowitz and Stegun, Chapter 22.

Dennery and Krzywicki, Chapter III, Section 10.

G. Szegő, "Orthogonal Polynomials" (whole book!).

A) Defining Property

$$\int_a^b w(x) p_n(x) p_m(x) dx = h_n \delta_{nm}$$

B) Recurrence Relation

$$p_{n+1}(x) = (A_n x + B_n) p_n(x) - C_n p_{n-1}(x)$$

Let  $k_n$  = coefficient of  $x^n$  in  $p_n(x)$

$k_{n-1}$  = coefficient of  $x^{n-1}$  in  $p_n(x)$

Then  $A_n = k_{n+1}/k_n$

$$B_n = \frac{k_{n+1}}{k_n} \left( \frac{k'_{n+1}}{k_{n+1}} - k'_n/k_n \right)$$

$$C_n = \frac{h_n}{h_{n-1}} k_{n+1} k_{n-1}/k_n^2$$

C) Christoffel-Darboux Formula

$$\text{Let } K_n(x, y) = \sum_{k=0}^n p_k(x) p_k(y) / h_k$$

Theorem:

$$K_n(x, y) = \frac{k_n}{k_{n+1}} \left( \frac{p_n(y)p_{n+1}(x) - p_n(x)p_{n+1}(y)}{x - y} \right)$$

Theorem: Among all polynomials  $v_n(x)$ , of unit norm, and degree  $n$ , i.e.,

$$\int_a^b w(x) |v_n|^2(x) dx = 1, \quad \text{the one that maximizes } |v_n(y)| \text{ where } y \text{ is an}$$

arbitrary complex point is

$$v_n(x) = e^{i\theta} K_n(y^*, x) / K_n(y^*, y)$$

where  $\theta$  is an arbitrary real number.D) Zeros of Orthogonal PolynomialsTheorem: The zeros of  $p_n(x)$  are real, distinct and all lie in the interior of  $[a, b]$ .Theorem: The zeros of  $p_n(x)$  and  $p_{n+1}(x)$  interlace.Theorem: Between any two zeros of  $p_n(x)$ , there is a zero of  $p_{n+1}(x)$ ,  $n > 0$ .Theorem: Let  $(\alpha, \beta)$  be any subinterval of the finite interval  $[a, b]$ . Then for sufficiently large  $n$ ,  $p_n(x)$  vanishes at least once in  $(\alpha, \beta)$ .

E) Approximation

Theorem:  $d = ||f(x) - v_n(x)||$  is minimized by  $v_n(x) = \sum_{k=0}^n \alpha_k p_k(x) / \sqrt{h_k}$

where  $\alpha_k = (p_k, f) / \sqrt{h_k}$

Here  $v_n(x)$  is an arbitrary polynomial of degree  $n$ .

$$(g, f) = \int_a^b w(x) g^*(x) f(x) dx$$

$$||g|| = (g, g)^{1/2}$$

F) Gaussian Integration

Theorem: If  $x_1 \dots x_n$  are the zeros of  $p_n(x)$  and  $v_{2n-1}(x)$  is any polynomial of degree  $2n-1$ . Then

$$\int_a^b w(x) v_{2n-1}(x) dx = \sum_{i=1}^n \lambda_i v_{2n-1}(x_i)$$

where  $\lambda_i = 1/K_n(x_i, x_i)$

Classical Orthogonal Polynomials

These have all above properties plus

G) Generalized Rodriguez Formula

$$P_n(x) = C_{(n)}(x) = \frac{1}{K_n w(x)} \frac{d^n}{dx^n} [w(x)s(x)^n]$$

where

- (i)  $C_{(1)}(x)$  is a first degree polynomial in  $x$ .
- (ii)  $s(x)$  is a polynomial in  $x$  of degree  $\leq 2$ .
- (iii)  $w(x)$  is real positive, integrable and satisfies boundary conditions:  $w(a)s(a) = w(b)s(b) = 0$ .

H) Differential Equation

$$\frac{d}{dx} [sw \frac{d}{dx} C_{(n)}] = -w(x) \lambda_n C_{(n)}(x)$$

$$\lambda_n = -n \left[ K_1 \frac{d}{dx} C_{(1)} + \frac{1}{2} (n-1) \frac{d^2 s}{dx^2} \right]$$

1) Examples of Classical Polynomials(i) Hermite Polynomials  $H_n(x)$ 

Standardization	$K_n = (-1)^n$
Constants	$k_n = 2^n, \quad k'_n = 0, \quad h_n = \sqrt{\pi} 2^n n!$
Rodriguez formula	$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2})$
Differential equation	$\frac{d^2}{dx^2} H_n(x) - 2x \frac{d}{dx} H_n(x) + 2n H_n(x) = 0$
Recurrence formula	$H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x)$

(ii) Laguerre Polynomials  $L_n^v(x)$ 

Standardization	$K_n = n!$
Constants	$k_n = \frac{(-1)^n}{n!}, \quad k'_n = -\left(\frac{n+v}{n}\right) k_n, \quad h_n = \frac{\Gamma(n+v+1)}{n!}$
Rodriguez formula	$L_n^v(x) = \frac{1}{n!} x^{-v} e^{-x} \frac{d^n}{dx^n} (e^{-x} x^{v+n})$
Differential equation	$x \frac{d^2}{dx^2} L_n^v(x) + (v+1-x) \frac{d}{dx} L_n^v(x) + n L_n^v(x) = 0$
Recurrence formula	$(n+1) L_{n+1}^v(x) = (2n+v+1-x) L_n^v(x) - (n+v) L_{n-1}^v(x)$

(iii) Jacobi Polynomials  $P_n^{(\nu, \mu)}(x)$ Standardization  $K_n = (-2)^n n!$ Constants\*  $k_n = 2^{-n} \binom{2n+\nu+\mu}{n}$ ,  $k'_n = \frac{n(\nu-\mu)}{2n+\nu+\mu} k_n$ 

$$h_n = \frac{2^{\nu+\mu+1} \Gamma(n+\nu+1) \Gamma(n+\mu+1)}{(2n+\nu+\mu+1)n! \Gamma(n+\nu+\mu+1)}$$

Rodriguez formula  $P_n^{(\nu, \mu)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\nu} (1+x)^{-\mu} \frac{d^n}{dx^n} [(1-x)^{\nu+n} (1+x)^{\mu+n}]$ Differential equation  $(1-x)^2 \frac{d^2}{dx^2} P_n^{(\nu, \mu)}(x) + [\mu - \nu - (\nu + \mu + 2)x] \frac{d}{dx} P_n^{(\nu, \mu)}(x) + n(n + \nu + \mu + 1) P_n^{(\nu, \mu)}(x) = 0$ Recurrence formula  $2(n+1)(n+\nu+\mu+1)(2n+\nu+\mu) P_{n+1}^{(\nu, \mu)}(x) = (2n+\nu+\mu+1) \{ (2n+\nu+\mu)(2n+\nu+\mu+2)x + \nu^2 - \mu^2 \} P_n^{(\nu, \mu)}(x) - 2(n+\nu)(n+\mu)(2n+\nu+\mu+2) P_{n-1}^{(\nu, \mu)}(x)$ \*The symbol  $\binom{y}{x}$  should be read

$$\binom{y}{x} = \frac{\Gamma(y+1)}{\Gamma(x+1)\Gamma(y-x+1)}$$

which for integer  $y$  and  $x$  reduces to the well-known expression

$$\frac{y!}{x!(y-x)!}$$

(iv) Gegenbauer Polynomials  $C_n^\lambda(x)$  (special case of (iii))

Standardization  $K_n = (-2)^n n! \frac{\Gamma(n+\lambda+1)\Gamma(2\lambda)}{\Gamma(n+2\lambda)\Gamma(\lambda+1/2)}$

Constants  $k_n = \frac{2^n \Gamma(n+\lambda)}{n! \Gamma(\lambda)}$ ,  $k'_n = 0$

$$h_n = \frac{\sqrt{\pi} \Gamma(n+2\lambda) \Gamma(\lambda+1/2)}{(n+\lambda)n! \Gamma(\lambda) \Gamma(2\lambda)}$$

Rodriguez formula  $C_n^\lambda(x) = \frac{(-1)^n \Gamma(n+2\lambda) \Gamma(\lambda+1/2)}{2^n n! \Gamma(n+\lambda+1/2) \Gamma(2\lambda)} (1-x^2)^{-\lambda+1/2} \frac{d^n}{dx^n} [(1-x^2)^{n+\lambda-1/2}]$

Differential equation  $(1-x^2) \frac{d^2}{dx^2} C_n^\lambda(x) - 2(\lambda+1)x \frac{d}{dx} C_n^\lambda(x) + n(n+2\lambda) C_n^\lambda(x) = 0$

Recurrence formula  $(n+1)C_{n+1}^\lambda(x) = 2(n+\lambda)x C_n^\lambda(x) - (n+2\lambda-1)C_{n-1}^\lambda(x)$

(v) Legendre Polynomials  $P_n(x)$  (special case of (iv))

Standardization  $K_n = (-2)^n n!$

Constants  $k_n = \frac{2^n \Gamma(n+1/2)}{n! \Gamma(1/2)}$ ,  $k'_n = 0$

$$h_n = (n+1/2)^{-1}$$

Rodriguez formula\*  $P_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} [(1-x^2)^n]$

---

\*The generalized Rodriguez formula is a generalization of this particular formula, originally derived by Rodriguez.

Differential equation  $(1-x^2) \frac{d^2}{dx^2} P_n(x) - 2x \frac{d}{dx} P_n(x) + n(n+1) P_n(x) = 0$

Recurrence formula  $(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$

(vi) Tchebichef Polynomials of the First Kind  $T_n(x)$  (special case of (iii))

Standardization  $K_n = (-1)^n \frac{(2n)!}{2^n n!}$

Constants  $k_n = 2^{n-1}$ ,  $k'_n = 0$ ,  $h_n = \frac{\pi}{2}$

Rodriguez formula  $T_n(x) = \frac{(-1)^n 2^n n!}{(2n)!} (1-x^2)^{-1/2} \frac{d^n}{dx^n} [(1-x^2)^{n-1/2}]$

Differential equation  $(1-x^2) \frac{d^2}{dx^2} T_n(x) - x \frac{d}{dx} T_n(x) + n^2 T_n(x) = 0$

Recurrence formula  $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$

(vii) Tchebichef Polynomials of the Second Kind  $U_n(x)$  (special case of (iii))

Standardization  $K_n = \frac{(-1)^n (2n+1)!}{2^n (n+1)!}$

Constants  $k_n = 2^n$ ,  $k'_n = 0$ ,  $h_n = \frac{\pi}{2}$

Rodriguez formula  $U_n(x) = \frac{(-1)^n 2^n (n+1)!}{(2n+1)!} (1-x^2)^{-1/2} \frac{d^n}{dx^n} [(1-x^2)^{n+1/2}]$

Differential equation  $(1-x^2) \frac{d^2}{dx^2} U_n(x) - 3x \frac{d}{dx} U_n(x) + n(n+2) U_n(x) = 0$

Recurrence formula  $U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x)$

B7 Fourier Series (M. and W. Section 4-2).

(i) So far we have considered the general problem of expanding functions  $f(x)$  (or abstractly vectors  $|f\rangle$ ) in terms of a complete basis  $e_n(x)$  ( $|e_n\rangle$ ) that is orthonormal.

$$\langle e_n | e_m \rangle = \delta_{nm}. \quad (B7.1)$$

Then we took the Special Hilbert Space  $L^2_w[a, b]$  and considered the special basis functions gotten by orthogonalizing the polynomials in  $x$ .

Now we consider the same ~~of~~ Space but a different type of expansion function; namely the trigonometric functions  $\cos nx$ ,  $\sin nx$ . Formally it is pretty clear, how we relate Fourier Series to our previous general formalism. Take  $w(x) = 1$ ,  $a = -\pi$  and

$l = \pi$ : then let the label  $n$  run from  $-\infty$  to  $+\infty$ , rather than 1 to  $\infty$ .

$$\text{Put } |e_n\rangle = \frac{1}{\sqrt{2\pi}} e^{in\theta} \quad (B7.2)$$

$$\text{and as } \int_{-\pi}^{\pi} \frac{(e^{in\theta})^*}{\sqrt{2\pi}} \frac{(e^{im\theta})}{\sqrt{2\pi}} d\theta = \delta_{nm},$$

this is certainly an orthonormal basis.

$$\text{we write } f(\theta) = \sum_{n=-\infty}^{\infty} a_n \frac{e^{in\theta}}{\sqrt{2\pi}} \quad (B7.3)$$

$$\text{where } a_n = \int_{-\pi}^{+\pi} f(\theta) \frac{e^{-in\theta}}{\sqrt{2\pi}} d\theta = \langle \hat{f}_n | f \rangle$$

and we have realised our standard formalism and obtained the Fourier Series. We do not repeat the standard stuff about Parseval's relation and Bessel's inequality. Rather we must ask is  $|e_n\rangle$  a complete set and if complete to which norms (i.e. in what space) is it complete. The latter is of course

equivalent to saying in what we mean the equality  $|f\rangle = \sum_{n=-\infty}^{\infty} a_n |e_n\rangle$  in (87.3). If we hadn't our previous experience we would have asked, no doubt, that (87.3) hold in the sense of either point convergence or (i.e.  $\sum_{n=-\infty}^{+\infty} a_n e^{in\theta} / \sqrt{2\pi} \rightarrow f(\theta)$  for each  $\theta$ ) or

even better uniform convergence. As we have explained many times this is a hard problem. So we first consider convergence in the mean i.e.

$$\int_{-\pi}^{\pi} \left( f(\theta) - \sum_{-n}^{+n} a_n \frac{e^{in\theta}}{\sqrt{2\pi}} \right)^2 d\theta \rightarrow 0$$

and only then consider point convergence

### (ii) Convergence in the Mean

We must quote an unproved theorem

(D. and K, top p.202, also just quote result).

## Generalized Weierstrass Theorem

If  $f(x_1 \dots x_s)$  is continuous in the domain  $a_i \leq x_i \leq b_i$  :  $(i=1, 2, \dots, s)$ . It can be uniformly approximated by ~~the~~ linear combinations of the monomials:  $x_1^{m_1} x_2^{m_2} \dots x_s^{m_s}$  :  $m_i \geq 0$ .

This, for  $s=1$ , reduces to Weierstrass's theorem on the approximation of any continuous function by polynomials. The proof is the obvious extension of that given on pages 50-52 for the case  $s=1$ . Thus you explicitly show that one can use the polynomials:

$$P_n(x_1, \dots, x_m) \approx \int_{a_1}^{b_1} \dots \int_{a_m}^{b_m} f(u_1, \dots, u_m)$$

$$\cdot \{ (1 - (u_1 - x_1)^2) (1 - (u_2 - x_2)^2) \dots (1 - (u_m - x_m)^2) \}^n du_1 \dots du_m$$

- See Courant and Hilbert - page 68.

So consider any function  $f(\theta)$  that is continuous in  $[-\pi, \pi]$  and is periodic i.e.

$$f(-\pi) = f(\pi). \quad (87.4)$$

Now we go into polar co-ordinates

$$x = r \cos \theta$$

$$y = r \sin \theta$$

where  $r$  is a new variable which we extend  $f(\theta)$  to be a function of by

$$f(\theta) \rightarrow r f(\theta) \equiv f(r, \theta)$$

Then this defines a function  $f(x, y)$  which is clearly continuous in the square  $-1 \leq x \leq 1, -1 \leq y \leq 1$ . Then by the previous  $f(x, y)$  can be uniformly approximated by a monomial

$$f_n(x, y) = \sum_{0 \leq m_i, n_i \leq n} x^{m_i} y^{n_i} a_{ij}^{(n)}$$

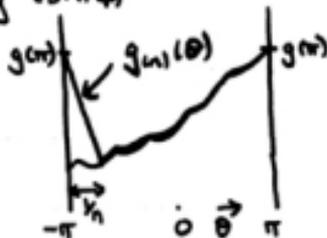
In particular it is uniformly approximated by  $f_n$  on the subset  $r=1$  of our square

domain :  Square domain. This gives

$f(\theta)$  uniformly approximated by  $\sum_{0 \leq m, n \leq n} a_{ij}^{(n)} \cos^m \theta \sin^n \theta$ , or substituting  $2 \cos \theta = e^{i\theta} + e^{-i\theta}$ ,  $2 \sin \theta = e^{i\theta} - e^{-i\theta}$ , we get

$f(\theta)$  uniformly approximated by  $f^{(n)}(\theta) = \sum_{m=-n}^n b_m^{(n)} e^{im\theta}$

where the  $b_m^{(n)}$  are linear combinations of  $a_{ij}^{(n)}$ . We can now relax the condition (B.7.4) as any continuous function  $g(\theta)$  can be considered as the limit of  $g_n(\theta)$  which satisfy (B.7.4)



$g_n(\theta) \rightarrow g(\theta)$  in the sense of convergence in the mean. It follows all continuous functions in  $[-\pi, \pi]$  can be arbitrarily well

approximated by linear combinations of  $e^{in\theta}$  i.e.  $\|g(\theta) - \sum_{m=-n}^n b_m^{(n)} e^{im\theta}\| < \epsilon$  for some  $n$ ,  $b_m^{(n)}$  and  $\|\cdot\|$  is usual norm in  $L^2_w[-\pi, \pi]$ . Thus it follows just as for the polynomials that  $e^{in\theta}$  ( $-\infty \leq m \leq \infty$ ) are complete in  $L^2_w[a, b]$  as this space was completion of continuous functions wrt  $\|\cdot\|$ .

### (iii) Pointwise Convergence

We now consider the harder problem of the limit, for a given  $\theta$ , as  $n \rightarrow \infty$  of the Fourier series (87.3). The nicest result is:

Theorem (D. and K. page 221, M. and W. pages 96-97).

The Fourier series of a function  $f(\theta)$  that is of bounded variation for  $-\pi \leq \theta \leq \pi$  converges

$$\text{to: } \frac{1}{2} [f(\theta_+) + f(\theta_-)] \quad \text{for } -\pi < \theta < \pi$$

$$\frac{1}{2} [f(\pi) + f(-\pi)] \quad \text{for } \theta = \pi.$$

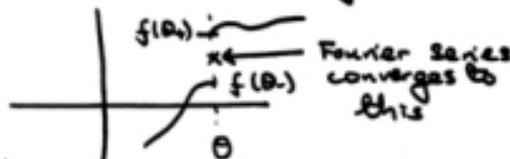
$f(\theta)$  is further convergent is uniform in any subinterval where

Now, by  $f(0_+)$  we mean

$$\lim_{\varphi \rightarrow 0 \text{ from above}} \{f(\varphi)\}$$

and  $f(0_-)$  is the corresponding limit from below.

At a point of continuity of  $f(x)$ ,



The picture shows  $f(0_+) = f(0_-)$  of course. ~~It's~~ a jump

discontinuity (functions of bounded variation can have ~~denumerably~~ <sup>finitely</sup> many of these and at their discontinuities  $f(0_{\pm})$  both exist, i.e. the picture shows the only possible allowed discontinuity).

Now we must define a function to be of bounded variation if  $\exists$  a bound  $B$ , such that the variation

$$V(x_1, \dots, x_n) < B \quad \text{for all choices of ordered sequences } x_1, \dots, x_n \text{ (any } n) \text{ in } [-\pi, \pi]. \quad (x_1 < x_2 < \dots < x_n).$$

$$\text{Here } V(x_1 \dots x_n) = |f(x_1) - f(x_2)| + |f(x_2) - f(x_3)| \\ + \dots + |f(x_{n-1}) - f(x_n)|.$$

Although this is a rather precise theorem, it may be considered useless as nobody has too much feeling as to whether or not a given function has bounded variation. Take an example:

Theorem: Any function with a continuous first derivative has bounded variation

Proof: It is well-known that

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(c) \quad : \quad x_1 \leq c \leq x_2$$

where the derivative  $f'(x)$  is continuous

in  $x_1 \leq x \leq x_2$ . Then

$$V(x_1 \dots x_n) = |x_1 - x_2| |f'(c_{1,2})| + \dots$$

$$\leq B' \{ |x_1 - x_2| + \dots + |x_{n-1} - x_n| \}$$

$$\leq B'(b-a) \quad \text{where the continuous}$$

$f(x)$  must be bounded and so it was legal to define  $B'$  so that

$$|f'(f)| < B' : \text{any } f \in [a, b]$$

In this special case - if we also add the periodicity condition, it is elementary to prove the main theorem on page 827

Theorem : Let the function  $f(\theta)$  and its derivative be continuous for  $-\pi \leq \theta \leq \pi$ , and let it satisfy the periodicity condition:

$$f(\pi) = f(-\pi)$$

Then the Fourier Series

$$\frac{1}{\sqrt{2\pi}} \sum_{-\infty}^{+\infty} a_n e^{in\theta} \quad (B7.3)$$

with  $a_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{+\pi} f(\theta) e^{-in\theta} d\theta$

converges uniformly to  $f(\theta)$  in the interval  $[-\pi, \pi]$

Proof: (D. and K. : page 219-220. They have a trivial error in proof of Eqn (11.15) but result is right).

as  $f'(\theta)$  is continuous, it certainly lies in  $L^2_1(-\pi, \pi)$  and so can be expanded

$$|f'(\theta)\rangle = \sum_{-\infty}^{+\infty} \langle e_n | f'(\theta) \rangle |e_n\rangle$$

which converges in the mean. Now let's

find  $\langle e_n | f'(\theta) \rangle$  in terms of  $a_n$ .

$$\begin{aligned} \langle e_n | f'(\theta) \rangle &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{+\pi} e^{-in\theta} f'(\theta) d\theta \\ \text{integrate by parts} &= \frac{1}{\sqrt{2\pi}} \left[ f(\theta) e^{in\theta} \right]_{-\pi}^{\pi} + \frac{in}{\sqrt{2\pi}} \int_{-\pi}^{+\pi} f(\theta) e^{-in\theta} d\theta \end{aligned}$$

$$= in a_n.$$

But Bessel's equality (Parseval's equation)

implies

$$\langle f'(\theta) | f'(\theta) \rangle = \sum_{n=-\infty}^{+\infty} |in a_n|^2 < \infty.$$

Now consider condition on  $a_n$  that

convergence of this series implies.

i.e. given  $\epsilon$ ,  $> k$

$$\sum_{|n| \geq k} |a_n|^2 < \epsilon \quad (87.5)$$

Now

$$\left| \sum_{k \leq |n| \leq K} a_n e^{in\theta} \right| \leq \sum_{k \leq |n| \leq K} |a_n|$$

$$\leq \sum_{k \leq |n| \leq K} |a_n| \cdot \frac{1}{|n|}$$

$$\leq \left( \sum_{k \leq |n| \leq K} |a_n|^2 \right)^{1/2} \left( \sum_{k \leq |n| \leq K} \frac{1}{n^2} \right)^{1/2}$$

$\leq \epsilon$  for large enough  $k$  as both series are convergent. So by Cauchy criterion, the series (87.3) is uniformly convergent and proof is complete

### Examples

In practice, we usually expand in terms of  $\cos n\theta$  and  $\sin n\theta$  rather than their (complex) linear combinations  $e^{\pm in\theta}$ .

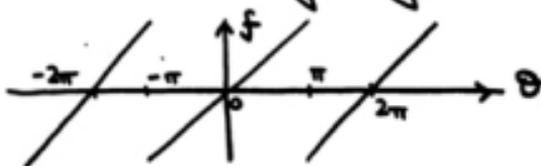
$\cos n\theta \pm i \sin n\theta$ . Then we rewrite (87.3) as

$$f(\theta) = \frac{A_0}{2} + \sum_{n=1}^{\infty} (A_n \cos n\theta + B_n \sin n\theta)$$

$$A_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\theta) \cos n\theta \, d\theta \quad (87.6)$$

$$B_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\theta) \sin n\theta \, d\theta$$

Take the function  $f(\theta) = \theta$ . This is defined for  $-\pi < \theta < \pi$ . We continue it outside this range by continuity

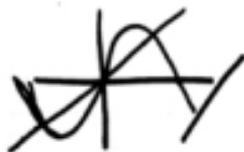


This is an odd function and so there are no cosine terms,  $A_n = 0$ .

$$\begin{aligned}
 B_n &= \frac{2}{\pi} \int_0^{\pi} \theta \sin n\theta \, d\theta \quad : \text{integrate by parts} \\
 &= \left[ -\frac{\cos n\theta}{n} \cdot \frac{2}{\pi} \theta \right]_0^{\pi} + \int_0^{\pi} \frac{2}{\pi n} \cos n\theta \, d\theta \\
 &= -(-1)^n \frac{2}{n}
 \end{aligned}$$

$$f(\theta) = 2 \left\{ \sin\theta - \frac{\sin 2\theta}{2} + \frac{\sin 3\theta}{3} \dots \right\}$$

The first approximation  $2\sin\theta$  looks like



: let's examine convergence in greater detail.

$$\text{Let } f_N(\theta) = 2 \sum_{n=2}^N \frac{(-1)^{n+1} \sin n\theta}{n}$$

$$= \frac{1}{i} \sum_{n=1}^N \left\{ \frac{(-1)^{n+1} e^{in\theta}}{n} - \frac{(-1)^{n+1} e^{-in\theta}}{n} \right\}$$

$$f'_N(\theta) = \sum_{n=1}^N \left\{ (-1)^{n+1} e^{in\theta} + (-1)^{n+1} e^{-in\theta} \right\}$$

$$= \left\{ \frac{e^{i\theta} (1 + e^{iN\theta} (-1)^{N+1})}{1 + e^{i\theta}} + \frac{e^{-i\theta} (1 + e^{-iN\theta} (-1)^{N+1})}{1 + e^{-i\theta}} \right\}$$

$$= \left\{ 1 + (-1)^{N+1} \left\{ \frac{\cos(N+1)\theta + \cos N\theta}{1 + \cos\theta} \right\} \right\}$$

Integrate using  $f_N(\theta) = 0$  at  $\theta = -\pi$

$$f_N(\theta) = \pi + \theta + (-1)^{N+1} \int_{-\pi}^{\theta} dt \frac{\cos(N+1/2)t}{\cos 1/2 t}$$

$$\text{or } f_N(\theta) - f(\theta) = \pi - \int_0^{\varphi} \frac{\sin(N+1/2)t}{\sin 1/2 t} dt$$

where  $\varphi = \theta + \pi$ .

Put  $(N+1/2)t = y$

$$f_N(\theta) - f(\theta) = \pi - \int_0^{(N+1/2)\varphi} \frac{\sin y}{\sin(y/(2N+1))} dy \frac{1}{(2N+1)}$$

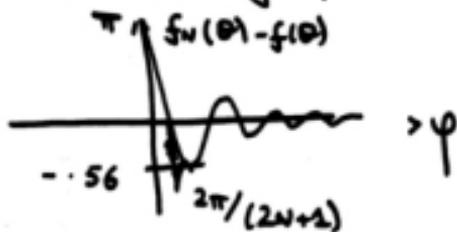
for  $\varphi$ , let  $N \rightarrow \infty$ , then by Staked convergence  $f_N(\theta) - f(\theta) \rightarrow 0$  if  $\theta \neq \pm\pi$

$1/2N+1/2$ . Small  $\psi$

Now for small  $\psi$ , we get

$$f_N(\psi) - f(\psi) = \pi - 2 \int_0^{(2N+1)\psi} \sin y / y \, dy$$

and using  $\int_0^{\infty} \sin y / y \, dy = \pi/2$ , we  
get this. little subtle as "oo" violates derivation  
But graph is



where maxima/minima occur when  
integrand  $\sin y = 0$  i.e.

$$\psi = 2\pi k / (2N+2)$$

The size decreases as  $k$  increases.

The maximum deviation is at

$$\psi_2 = 2\pi / (2N+2)$$

when

$$f_N(\psi) - f(\psi) = \pi - 2 \int_0^{\pi} \sin y / y \, dy \approx -.56$$

a value independent of  $N$ .

This doesn't contradict convergence as position  $\varphi_2 \rightarrow 0$  as  $N \rightarrow \infty$



It is called Gibbs' phenomena and illustrates peculiar lack of uniform convergence near discontinuity of  $f(\theta)$ .

88 Fourier Integrals (read m. and w. sections 4-2 to 4-5)

So far we have only considered functions with period  $2\pi$ :

$$f(\theta + 2\pi) = f(\theta).$$

The use of fourier series for a function of period  $2L$  is trivial:

$$\text{if } f(x + 2L) = f(x)$$

Put  $\theta = \pi x/L$  and regarded as a function of  $\theta$ ,  $f$  has period  $2\pi$ . (87.3)

becomes:

$$f(x) = \frac{1}{\sqrt{2L}} \sum_{-\infty}^{\infty} a_n e^{in(\pi/L)x} \tag{88.1}$$

$$a_n = \frac{1}{\sqrt{2L}} \int_{-L}^{+L} f(x) e^{-in(\pi/L)x} dx$$

after renormalizing  $a_n$  to make it more symmetric.

Now put  $t = n\pi/L$  and  $\delta t = \pi/L$ .

For large  $L$ , we can replace Sum over  $n$

by an integral  $dt$ . (as increment  $n \rightarrow n+1$  corresponds to  $t \rightarrow t+\delta t$  and  $\delta t \rightarrow 0$  as  $L \rightarrow \infty$  for fixed  $t$ ). Again defining normalization of  $a_n$  by  $b_n = \sqrt{\frac{L}{\pi}} a_n$ , (88.1) becomes:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} b(t) e^{itx} dt$$

$$b(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-itx} dx \quad (88.2)$$

This is a Fourier integral which together with its almost trivially related brothers - the Mellin transform and the Laplace transform is used especially to help solve differential and integral equations. We will return to these applications later - now we just concentrate on some formal matters. Titchmarsh (Introduction to the theory of Fourier Integrals) has lots of

unreadable theorems about the niceties of the convergence of (88.2). A simple theorem, given without proof in both D. and K. and M. and W) is due to Plancherel

Theorem: Let  $|f(\theta)|^2$  be integrable in the interval  $(-\infty, \infty)$ . The integral

$$F(t, \Lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\Lambda}^{+\Lambda} f(\theta) e^{-i\theta t} d\theta \quad (88.3a)$$

converges in the mean. (remember what this is? - see page 36), as  $\Lambda \rightarrow \infty$ , to a function  $F(t)$  whose square modulus  $|F(t)|^2$  is also integrable in  $(-\infty, \infty)$ . Furthermore, the integral

$$f(\theta, \Lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\Lambda}^{+\Lambda} F(t) e^{i\theta t} dt \quad (88.3b)$$

converges in the mean to  $f(\theta)$  and

$$\int_{-\infty}^{+\infty} |f(\theta)|^2 = \int_{-\infty}^{+\infty} |F(t)|^2 \quad (88.4)$$

Note the nice symmetry between  $f(\theta)$  and  $F(t)$  in this theorem. Further (88.4) is

Parseval's relation

$$\langle x|x \rangle = \sum_{j=1}^{\infty} |\langle e_j|x \rangle|^2 \tag{B8.4'}$$

is modified in an obvious way as we converted sum into an integral at the beginning of this section. Now the

easy way to prove (B8.4') is to put

$$|x \rangle = \sum_{j=1}^{\infty} \langle e_j|x \rangle |e_j \rangle \tag{B8.5a}$$

$$\text{and use } \langle e_j|e_k \rangle = \delta_{jk}. \tag{B8.5b}$$

It is tempting to write (B8.2) for

some  $|x \rangle$  (e.g.  $\in L^2[-\infty, \infty]$ ) as

$$|x \rangle = \int_{-\infty}^{+\infty} \langle e(t)|x \rangle |e(t) \rangle dt \tag{B8.6a}$$

but what do we use in place of (B8.5b)?

Go ahead anyway:

$$\langle x|x \rangle = \int_{-\infty}^{\infty} dt \langle e(t)|x \rangle \int_{-\infty}^{\infty} dt' \langle e(t')|x \rangle^* \langle e(t')|e(t) \rangle \tag{B8.7}$$

$$\text{Define } \delta(t'-t) = \langle e(t')|e(t) \rangle \tag{B8.6b}$$

To achieve (88.4) from (88.7) it must

satisfy:

$$\int_{-\infty}^{+\infty} \langle e(t) | x \rangle^* \delta(t'-t) dt' = \langle e(t) | x \rangle^* \quad (88.8)$$

or realizing that  $\langle e(t) | x \rangle$  is essentially the arbitrary (integrable) function, we must have:

$$f(x) = \int_{-\infty}^{+\infty} f(y) \delta(x-y) dy \quad (88.9)$$

which is the fundamental property of the Dirac delta-function. This is clearly no ordinary function for (88.9) implies

$$\delta(x) = 0, \quad x \neq 0$$

$$\int_{-a}^b \delta(x) dx = 1 \quad a, b > 0. \quad (88.10)$$

If we are gay about interchanging limits we can find the Fourier transform of  $\delta(x)$  from the reciprocal relations (88.3a,b) or rather (88.2).

$$\begin{aligned}
 f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dt e^{tx} \mathcal{L}(f) \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dt e^{tx} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dy f(y) e^{-ty} dy \\
 &= \int_{-\infty}^{+\infty} dy f(y) \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{t(x-y)} dt \quad (88.11)
 \end{aligned}$$

Comparing (88.11) with (88.9), we find

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{tx} dt \quad (88.12)$$

i.e.  $\delta(x)$  has Fourier coefficients  $\frac{1}{\sqrt{2\pi}}$ . Note the dubious behaviour of  $\delta(x)$  as a function of  $x$  is reflected in bad behaviour of its Fourier coefficients as  $t \rightarrow \infty$  i.e. the integral (88.12) doesn't really converge with normal definition of integrals. Now the theory of distributions was started in this hazy way by Dirac in 1928 and it was not until 1951 that Schwarz made it rigorous. The

mathematics is quite neat and so we will sketch it. let us summarize some difficulties

(i) Derivation of correct formal properties e.g. (88.4) requires use of  $\delta$ -function which is not a function. Further its Fourier series diverges according to normal rules

(ii) The latter difficulty is also seen if, say, I take  $f(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega x} b(t) dt$

where  $b(t) \sim 1/t^2$  as  $t \rightarrow \infty$ . Then the above is convergent and well-defined.

$$f'(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega x} (it b(t)) dt$$

has in general a Fourier expansion which is closely related to that of  $f(\omega)$ . of course this easy relation is why Fourier and Laplace transforms are used to facilitate

solution of D.E.'s. Unfortunately in our case, the integral for  $f'(0)$  is no longer convergent and so we don't seem to be able use formal relation.

- (iii) The reader may have noticed that it is little odd that in Plancherel's theorem, we may were expanding a function  $\in L^2[-\infty, \infty]$  in terms of a continuous set  $e^{itx}$  of functions; meanwhile in 86 we have used a denumerable set (Hermite functions) to expand the same functions.

These are all cleared up by both enlarging our space and redefining the meaning of equality. Note that this is a typical mathematical trick. Rather than answer a difficult question, you

change the question: (e.g. faced with the problem of finding the general criterion for the pointwise convergence of some expansion, you don't answer but show that convergence in mean allows an easy general statement).

## 89 Theory of Distributions

There are two approaches. The most high level mathematically is given in "Gelfand and Shilov" who have - with help from other authors - fully five volumes on "generalized functions" (a synonym for distributions). Volume 1 is elementary - its length coming from completeness (long-windedness?) not scope. Volume 2 covers abstract approach using the linear space formalism we introduced earlier. The other 3 volumes are applications which are outside the scope of the lectures and the comprehension of the lecturer. This method certainly allows the clearest understanding <sup>and why</sup> <sub>how</sub> distributions exist. However the next discussion is that of Lighthill.

("Fourier Analysis and Generalized Functions")  
 which uses a simpler but less general  
 method - D. and K. ~~for~~ <sup>precis</sup> Lighthill.  
 So we are forced to use this approach  
 but at the end we will indicate how  
 the more general methods work.

### (1) Good and Fairly Good Functions

Define: A good function is one which is  
 everywhere differentiable any number  
 of times and such that it and all its  
 derivatives are  $O(|x|^{-N})$  as  $|x| \rightarrow \infty$  for all  
 $N$ .

$e^{-x^2}$  is a good function as its  
 derivatives are (essentially) Hermite  
 polynomials  $\times e^{-x^2}$ . In the smooth  
 approach, good functions are an example  
 of a set of test functions. Note the set.

of good functions do form a linear vector space as they are closed on addition and multiplication by a constant.

Define: A sequence  $\{f_n(x)\}$  of good functions is called regular if, for any good function  $F(x)$  whatever, the limit

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} f_n(x) F(x) dx \text{ exists. (89.2)}$$

The two regular sequences  $\{f_n(x)\}$ ,  $\{g_n(x)\}$  are called equivalent if for any good function  $F(x)$ , the limit (89.2) is the same for the two sequences.

For example  $f_n(x) = e^{-x^2/n^2}$  is regular and equivalent to  $e^{-x^2/n^4}$ .

In this case the limit (89.2) is just

$$\int_{-\infty}^{+\infty} F(x) dx.$$

Define: A generalized function  $f(x)$  is defined as a regular sequence  $\{f_n(x)\}$  of good functions. But two generalized functions are said to be equal if the corresponding regular sequences are equivalent.

The integral  $\int_{-\infty}^{+\infty} f(x) F(x) dx$

of the product of a generalized function  $f(x)$  and a good function  $F(x)$  is defined as

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} f_n(x) F(x) dx \quad (B9.2)$$

which is unique as (B9.2) is the same for equivalent sequences.

Examples:

a) For our old example  $f_n(x) = e^{-x^2/n^2}$ , the associated generalized function is called  $I(x)$  - or  $\delta$  for short - and satisfies

$$\int_{-\infty}^{+\infty} I(x) F(x) dx = \int_{-\infty}^{+\infty} f(x) dx.$$

Generally given any integer ordinary function  $f(x)$  satisfying some weak integrability condition e.g.

$|f(x)| / (1+x^2)^N$ , any  $N \geq 0$  <sup>integer</sup> (B9.3) is integrable in  $[-\infty, +\infty]$

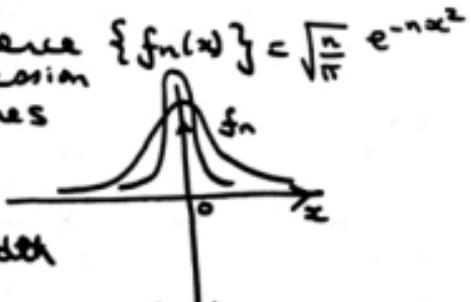
$\exists$  an associated generalization function  $\tilde{f}(x)$  satisfying

$$\int_{-\infty}^{+\infty} \tilde{f}(x) F(x) dx = \int_{-\infty}^{+\infty} f(x) F(x) dx.$$

So the class of generalized functions includes all sensible ordinary functions. However it includes many more.

(b) The Delta function

The sequence  $\{f_n(x)\}$  represents a succession of increasing spikes whose height increases and width



decreases as  $n \rightarrow \infty$ , in such a way as to

ensure  $\int_{-\infty}^{+\infty} f_n(x) dx = 1$ . The sequence

is regular - as is intuitively obvious -

for if  $F(x)$  is any good function:

$$\begin{aligned} & \left| \int_{-\infty}^{+\infty} e^{-nx^2} \sqrt{\frac{n}{\pi}} F(x) dx - F(0) \right| \\ &= \left| \int_{-\infty}^{+\infty} e^{-nx^2} \sqrt{\frac{n}{\pi}} (F(x) - F(0)) dx \right| \\ &\leq \max(|F'(x)|) \int_{-\infty}^{+\infty} e^{-nx^2} \sqrt{\frac{n}{\pi}} |x| dx \\ &= (\pi n)^{-1/2} \max(|F'(x)|) \text{ which } \rightarrow 0 \text{ as } \\ & n \rightarrow \infty. \end{aligned}$$

So  $\{f_n(x)\}$  defines a generalized function - called  $\delta(x)$  - with the property that

$$\int_{-\infty}^{+\infty} \delta(x) F(x) dx = F(0) \quad (89.4)$$

Note  $\delta$  is simply not defined as a function with values for each  $x$  - so

we avoid difficulties we found in (B2).  
 we can make sense of  $\delta(x) = 0$  for  $x \neq 0$ .  
 Thus define: If one of the regular  
 sequences, defining a generalized  
 function, converges uniformly to an  
 ordinary function in some neighbourhood  
 of  $x = x_0$ , then  $x = x_0$  is called a  
 regular point of the generalized function  
 and the limit of the sequence is called  
 the local value of the generalized  
 function at this point.

Being a uniform limit of continuous  
 functions the local values give a  
 continuous function in the neighbourhood  
 of  $x_0$  and it is trivial to prove that  
 it is unique. [any two sequences -  
 if they converge uniformly - give the same

local function]. So in this sense  $\delta(x) = 0, x \neq 0$  and  $I(x)$  (defined by  $e^{-x^2/2} = 1$ ).

### Properties of Generalized Functions

#### (i) Addition

If generalized functions  $f(x)$  and  $g(x)$  are defined by sequences  $\{f_n(x)\}$  and  $\{g_n(x)\}$ , then the sequence  $\{f_n(x) + g_n(x)\}$  is also regular and we define  $f(x) + g(x)$  to be the associated generalized function. This definition satisfies obvious consistency conditions.

#### (ii) Multiplication

The main restriction on generalized functions is that, in general,  $f(x)g(x)$  has no meaning. This comes from misfortune that if  $\{f_n(x)\}, \{g_n(x)\}$ ,

are regular,  $\{f_n \otimes g_n(x)\}$  isn't usually.

For instance if  $f_n(x) = g_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2}$

$f_n g_n = \frac{n}{\sqrt{\pi}} e^{-2nx^2}$  and

$$\int_{-\infty}^{+\infty} f_n(x) g_n(x) F(x) dx \propto \sqrt{n} F(0)$$

and has no limit.

However if we define a fairly good function to be one that is everywhere differentiable and such that it and its derivatives are  $O(|x|^{-N})$  for some  $N$ . Then the product of a good function and a fairly good function is still a good function. So

$$\int_{-\infty}^{+\infty} \underset{\substack{\uparrow \\ \text{fairly good}}}{p(x)} \underset{\substack{\uparrow \\ \text{regular}}}{f_n(x)} \underset{\substack{\uparrow \\ \text{good}}}{F(x)} = \int_{-\infty}^{+\infty} f_n(x) \underset{\substack{\uparrow \\ \text{good}}}{(p(x) F(x))}$$

has a well defined limit for each  $F(x)$  and fixed  $p(x)$ . Thus we may

define the product of any generalized function with a fairly good function. For instance  $x\delta(x)$  is defined and equals the generalized function 0. However  $(\delta(x))^2$  is not defined.

### (iii) Differentiation

Generalized functions are always differentiable! Thus if  $\{f_n(x)\}$  is regular we define the derivative of  $f(x)$  to be defined given by the regular

sequence  $\{df_n/dx\}$

$$\text{as } \int_{-\infty}^{+\infty} df_n/dx \cdot F(x) dx = - \int f_n(x) \frac{dF}{dx} dx$$

and  $dF/dx$  is always a good function, it is clear that  $\{df_n/dx\}$  is indeed a regular sequence.

As an example,  $\delta'(x)$  exists and satisfies

$$\int_{-\infty}^{+\infty} \delta'(x) F(x) dx = -F'(0)$$

Again let  $\theta(x)$  be the generalized function = 0 for  $x \leq 0$  and 1 for  $x > 0$ .

more precisely

$$\int_{-\infty}^{+\infty} \theta(x) F(x) dx = \int_0^{\infty} F(x) dx$$

Then from the definition:

$$\begin{aligned} \int_{-\infty}^{+\infty} \theta'(x) F(x) dx &= - \int_0^{\infty} F'(x) dx \\ &= [F(x)]_0^{\infty} = F(0) \end{aligned}$$

i.e.  $\theta'(x) = \delta(x)$

### (iv) Change of Argument

Let  $y(x)$  be any reasonable function - what do we mean by  $\delta(y(x))$  e.g.  $\delta(x-x_0)$  or  $\delta(x^2)$ .

All that counts is the behaviour of  $y(x)$  near its zeros as  $\delta$  is otherwise zero.

So let  $y(x_i) = 0 \quad i=1 \dots S$ . Then

$$\begin{aligned} I &= \int_{-\infty}^{+\infty} \delta(y(x)) F(x) dx \\ &= \sum_{i=1}^S \int_{x_i-\epsilon}^{x_i+\epsilon} \delta(y(x_i) + y'(x_i)(x-x_i)) F(x) dx \end{aligned}$$

Take each term separately and change variables:  $z = y'(x_i)(x-x_i)$

$$\begin{aligned} I &= \sum_{i=1}^S \int_{-\epsilon}^{+\epsilon} \delta(z) F(x_i \dots) dz / |y'(x_i)| \\ &= \sum_{i=1}^S F(x_i) / |y'(x_i)| \end{aligned}$$

$$\text{or } \delta(y(x)) = \sum_{i=1}^S \delta(x-x_i) / |y'(x_i)|$$

Note if  $y'(x_i) = 0$ , the  $\delta(y(x))$  is simply not defined. Take, for example  $\delta(x^3)$ , then as  $\delta(x)$  is limit of  $\sqrt{\frac{n}{\pi}} e^{-nx^2}$

So  $\delta(x^3)$  must be limit of

$$\sqrt{\frac{n}{\pi}} e^{-nx^6} \quad ; \text{ but this isn't}$$

a regular sequence for consider

$$\int_{-\infty}^{+\infty} \sqrt{\frac{n}{\pi}} e^{-nx^6} \cdot f(x) dx$$

where  $f(x) \sim 1$  for small  $x$ .

Then  $e^{-nx^6}$  is big for  $x \sim 1/n^{1/6}$

and contribution to integral is

$$\propto \sqrt{n} \times 1/n^{1/6} \rightarrow \infty \text{ as } n \rightarrow \infty$$

for  $\delta(x)$  we would get

$$\propto \sqrt{n} \times 1/\sqrt{n} \rightarrow \text{finite as } n \rightarrow \infty$$

### Fourier Transforms

This is what originally gave us an excuse to study generalized functions.

First note that if  $f(x)$  is a good function, its Fourier transform  $g(t)$  is also a good function. Thus

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-itx} dx$$

$$g^{(p)}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (ix)^p f(x) e^{-itx} dx$$

as  $x^p f(x)$  still  $\rightarrow 0$  fast at  $\infty$  and so differentiation justified. Integrate by parts  $N$  times

$$g^{(p)}(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{i^N} \int_{-\infty}^{+\infty} \frac{d^N}{dx^N} [ix^p f(x)] e^{-itx} dx$$

$$|g^{(p)}(t)| \leq \frac{1}{\sqrt{2\pi}} \frac{1}{|t|^N} \int_{-\infty}^{+\infty} \left| \frac{d^N}{dx^N} [ix^p f(x)] \right| dx$$

is  $O(t^{-N})$  any  $p, N$

which is the desired condition.

Now the theory of Fourier integrals for good functions is particularly simple: Plancherel's theorem (section 89) holds and in fact one can prove this and Parseval's relation

$$\int_{-\infty}^{+\infty} g_1(t)^* g_2(t) dt = \int_{-\infty}^{+\infty} f_1(x)^* f_2(x) dx \quad (89.5)$$

by elementary methods. (D. and K. p. 233).

Assuming this we define the Fourier integral transform of a generalized function by definition: given a regular sequence  $\{f_n(x)\}$  defining generalized function  $f(x)$ , then if  $\{g_n(x)\}$  are the F.T. of  $\{f_n(x)\}$ , then they also a regular sequence and define another generalized function  $g(x)$  called the F.T. of  $f(x)$ .

lets prove the regularity:

(89.5) implies

$$\int_{-\infty}^{+\infty} g_n(t) g(t) dt = \int_{-\infty}^{+\infty} f_n(x) F(x) dx \quad (89.6)$$

where  $g(t)$  is any good function and  $F^*(-x)$  is the F.T. of  $g^*(t)$ : by our theorem above  $F(x)$  is ~~also~~ a good function. So as  $\{f_n\}$  regular, the right hand side of (89.6) always exists in the limit as  $n \rightarrow \infty$ : thus (89.6)  $\Rightarrow$   $\{g_n(t)\}$  regular as  $g(t)$  was any good function.

The F.T., thus defined, as shows properties wrt differentiation, addition etc. Also if  $g(t)$  F.T. of  $f(x)$ ,  $f(-x)$  is F.T. of  $g(t)$ . Parseval's relation is valid as long as products of generalized functions well defined.

lets do an easy example

$$f(x) = x \delta'(x)$$

$$g(t) = \int_{-\infty}^{+\infty} e^{-itx} x \delta'(x) / \sqrt{2\pi} dx$$

$$= - \int_{-\infty}^{+\infty} \frac{d}{dx} (x e^{-itx}) \delta(x) dx$$

by integrating by parts

$$= - \left[ itx e^{-itx} + e^{-itx} \right] \Big|_{x=0} / \sqrt{2\pi}$$

$$= -1 / \sqrt{2\pi}$$

$$= - \text{F.T. of } \delta(x)$$

$$\therefore \underline{x \delta'(x) = -\delta(x)}$$

Let's consider a hard example:

$$f(x) = \sum_{m=-\infty}^{+\infty} \delta(x-m)$$

is a generalized function. If  $g_n(x) = \sqrt{n/\pi} \exp(-nx^2)$  is a sequence of good functions defining  $\delta(x)$ , then we may define  $f(x)$  by the sequence

$$e^{-x^2/n^2} \sum_{m=-\infty}^{+\infty} g_n(x-m)$$

Stoopingly we replace this by  $e^{-|x|/n} \sum_{m=-\infty}^{+\infty} \delta(x-m)$  which has essential features although it is not a good function. Its F.T. is elementary

$$\begin{aligned} \tilde{f}_n(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f_n(x) e^{-itx} dx \\ &= \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{+\infty} y^{|m|} e^{-itm} \end{aligned}$$

where  $y = e^{-1/n}$  is  $< 1$ .

$(\sqrt{2\pi} \widehat{f}_n(t) + 1)$  is sum of two geometric series

$$1 + z_1 + z_1^2 + \dots \quad z_1 = e^{ity}$$

$$+ 1 + z_2 + z_2^2 + \dots \quad z_2 = e^{-ity}$$

$$= \frac{1}{1-z_1} + \frac{1}{1-z_2}$$

$$= \frac{2 - 2\cos ty}{1 + y^2 - 2\cos ty}$$

Take limit  $y \rightarrow 1$  : if  $\cos t \neq 1$ , this is safe and we get 1 on r.h.s.

$$\text{i.e. } \widehat{f}_n(t) = 0 \quad \cos t \neq 1$$

Now take  $\cos t = 1$  i.e.  $t = 2m\pi$ .

For example if  $t$  small  $\approx 0$

$$\sqrt{2\pi} \widehat{f}_n(t) = \frac{1 - y^2}{1 + y^2 - 2y + t^2 y} \quad \text{using } \cos t = 1 - \frac{t^2}{2}$$

$$= \frac{2\epsilon}{\epsilon^2 + t^2} \quad \text{where } \epsilon = 1 - y \text{ is small}$$

It is easy to show

$$\lim_{\epsilon \rightarrow 0} \frac{\epsilon}{\epsilon^2 + t^2} = \pi \delta(t)$$

e.g.  $\int_{-\infty}^{+\infty} \frac{\epsilon dt}{\epsilon^2 + t^2} \quad v = t/\epsilon$

$$= \int_{-\infty}^{+\infty} \frac{dv}{1+v^2} = \left[ \tan^{-1} v \right]_{-\infty}^{+\infty} = \pi$$

and  $\lim_{\substack{\epsilon \rightarrow 0 \\ t \neq 0}} \frac{\epsilon}{\epsilon^2 + t^2} = 0$

Using this we get

$$\lim_{n \rightarrow \infty} \tilde{f}_n(t) = \sqrt{2\pi} \sum_{m=-\infty}^{m+\infty} \delta(t - 2m\pi)$$

which is then F.T. of  $\sum_{m=-\infty}^{m+\infty} \delta(x - m)$

## Mathematical Approach

Let's sketch the formal approach to generalized functions. As in simple method, we introduce the vector of good or test functions - call it  $K$ . Then we introduced a notion of convergence in  $K$ . We say a sequence of functions  $f_n(x) \rightarrow f$  in  $K$  if  $g(x) = f(x) - f_n(x) \rightarrow 0$  and all its derivatives  $\rightarrow 0$  uniformly in  $x$ . [There are some niceties here - Gelfand ~~now~~ <sup>first</sup> ~~considers~~ considers sequences which vanish for  $|x| > a$ : some  $a$ ; I'm not too sure if you read this intermediate step]. This definition of convergence is more general than those we saw before: in particular you cannot define a norm on this

Space which gives this convergence criterion. This is because convergence <sup>/boundedness</sup> was defined to be uniform in  $x$  and not in  $p$  (order of differentiation) and obvious norm  $\max_{\substack{x \in [-a, a] \\ 0 \leq p \leq \infty}} |g^{(p)}(x)|$  is not necessarily finite. Rather Gelfand uses notion of countably normed space.

Anyhow given this vector space, we consider functionals on it. These are a law  $\alpha$  which associates with any  $|f\rangle \in \text{Space}$  a number in  $-\infty \leq x \leq \infty$  (or a complex number). Generalized functions are continuous linear functionals.

$$\alpha(\alpha|f\rangle + \beta|g\rangle) = \alpha\alpha(|f\rangle) + \beta\alpha(|g\rangle)$$

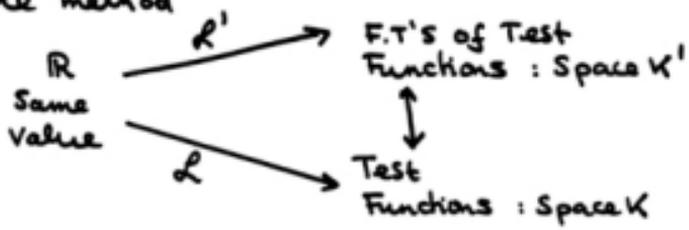
$$\alpha(|f_n\rangle) \rightarrow \alpha(|f\rangle) \text{ if } |f_n\rangle \rightarrow |f\rangle$$

Notice our old definition  $\int_{-\infty}^{+\infty} \delta(x) f(x) dx$

satisfies this although the continuity is tricky to prove for a true generalized function  $\delta(x)$  (limit  $\int_{-\infty}^{+\infty} \delta_n(x) f(x) dx$ ).

So functions map linear space  $\mathbb{R}$  onto  $\mathbb{C}$  : functionals map more general linear space  $K$  onto  $\mathbb{C}$ .

Fourier transforms are defined as in simple method



We showed  $K$  and  $K'$  were isomorphic spaces, and introduce functionals on both. We define F.T.  $L'$  by of  $L$  by

$$L'(f') = L(f) \text{ where } f' \text{ is}$$

F.T. of  $f$ . Parsevals relation (89.5)  
shows this definition is consistent.

## B10 Other Integral Transforms and Examples

(i) We showed in the last section how to make use of the divergent Fourier transform of say  $f(x) = c$ , a constant: it was  $c/\sqrt{2\pi} \delta(t)$ . Another way which is more useful in solving D.E.'s is by analytic continuation in the argument  $t$ . Suppose  $f(x)$  vanishes for  $x < 0$  or if it doesn't, we are not interested in its values for  $x < 0$  and can replace it by  $f(x) \theta(x)$  {remember  $\theta$ , the theta or Heaviside function = 1 for  $x > 0$  and 0 for  $x < 0$ }. Then its F.T.

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(x) e^{-itx} dx \quad (\text{B10.1})$$

Now if  $f(x)$  is bounded by some exponential ( $f(x)$  is  $O(e^{cx})$  - not satisfied by  $e^{x^2}$  for instance), then (B10.1) is

well-defined and indeed defines an analytic function of  $t$  for  $\text{Im } t < -c$ .

Let  $c'$  be any number  $> c$  and put

$$h(x) = f(x) \exp(-c'x)$$

~~$t = -c + t'$~~   $t = -c' + t'$

$$g(t) = \int_0^{\infty} h(x) e^{-t'x} dx / \sqrt{2\pi}$$

where now  $t'$  is real and  $h(x)$

decreases exponentially at  $\infty$ . Plancherel's theorem is applicable so that

$$h(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{it'x} dx \quad (8.10.2)$$



we can verify from (10.2) that

$g(t)$  is not only analytic for  $\text{Im } t' < 0$  but also behaves nicely as  $t' \rightarrow \infty$  under this condition. It follows from

(B10.2) that  $h(x)$  does indeed  $= 0$  for  $x < 0$ .  
 as we verify by using Cauchy's theorem to  
 complete integral by closed contour at  $\infty$ .

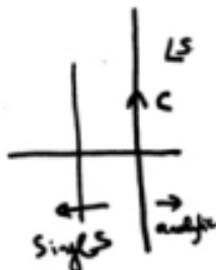
This is called a Laplace transform

Putting  $s = \sigma + it$  we get

$$L(s) = \int_0^{\infty} f(x) e^{-sx} dx$$

(B10.2)

$$f(x)\theta(x) = \frac{1}{2\pi i} \int_C L(s) e^{sx} ds$$



where  $C$  is contour  $\text{Re}(s) = c' > c$   
 and  $f(x)$  is  $O(e^{cx})$ .

As an example if  $f(x) = 1$ ,  $L(s) = 1/s$   
 and it is indeed analytic for  $\text{Re } s > 0$ .  
 There is a pole at  $s = 0$  with  $\text{Res} = c = 0$ .  
 Note that F.T. of 1 was  $\frac{1}{\sqrt{2\pi}} \delta(x)$  but  
 really we should compare with F.T of  
 $\theta(x)$  which is  $\frac{1}{\sqrt{2\pi}} \frac{1}{+it}$  ~~and there~~  
~~is an interesting sign discrepancy. (One~~  
 which agrees with Laplace Transform

after supplying the conventional  $1/\sqrt{2\pi}$ .  
~~would have expected  $1/\sqrt{2\pi}$   $1/t$  using~~  
 ~~$s \rightarrow t$  and supplying conventional  $1/\sqrt{2\pi}$ .~~

(ii) Similar to Laplace transform is Mellin transform:

$$\phi(z) = \int_0^{\infty} t^{z-1} f(t) dt$$

$$f(t) = \frac{1}{2\pi i} \int_C t^{-z} \phi(z) dz \quad : C \text{ is suitable } \text{Re } z = \text{const.}$$

which is derived from (B10.2) by

$$\log t = \int t^z dt/t = e^{-xz} = xz dx$$

This is used a lot in high energy physics as "Regge theory" claims  $\phi(z)$  should consist of poles in the  $z$ -plane when we take  $f(t)$  as scattering amplitude.

(iii) Also used in high-energy physics is the Fourier-Bessel transform where  $f(x)$  is scattering ampl as a function of  $x = \sqrt{-t}$  etc :  $t$  is invariant

momentum transfer and  $g(b)$  is partial wave amplitude as a function of impact parameter  $b$ .

$$g(b) = \int_0^{\infty} f(x) J_n(bx) x dx \quad (B10.4)$$

$$f(x) = \int_0^{\infty} g(b) J_n(bx) b db$$

(iv) Hilbert transforms are

$$g(y) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{f(x) dx}{x-y} \quad (B10.5)$$

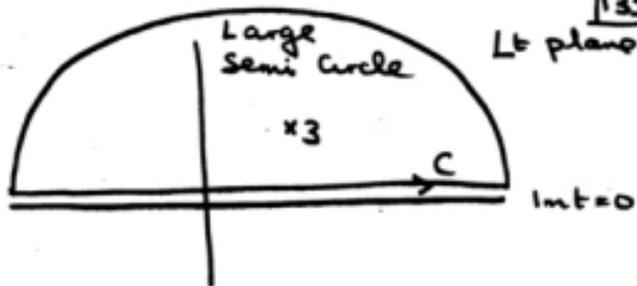
$$f(x) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{g(y) dy}{y-x}$$

These are valid under quite weak conditions for  $f$  and  $g$ . However they are most familiar when  $f$  and  $g$  are the real and imaginary parts of a function of a complex variable that satisfies some suitable analyticity and asymptotic conditions.

Then (B10.5) are called dispersion relations and they are discussed in M. and W. p. 129 to 135. We can sketch them here as they have some connection with the considerations we went through for Laplace transforms:

### (v) Dispersion Relations

In the Laplace transform section, we showed that F.T.'s of functions that were ~~on~~ zero for  $x < 0$  were analytic for  $\text{Im } t < -c$ . Now in high energy physics, the scattering amplitudes  $f_{\pm}(t)$  are such Fourier transforms (except  $t \rightarrow -t$ !) and further ~~in some~~ ~~circumstances~~  $c = 0$  and  $f_{\pm}(t) \rightarrow 0$  as  $t \rightarrow \infty$ . Consider any point  $z$  with  $\text{Im } z > 0$  and the contour  $C$  shown



$$I = \frac{1}{2\pi i} \int_C f(t) / (t-z) dt$$

evaluate  $I$  by Cauchy's theorem  
to get  $I = f(z)$  : the residue at  $t=z$ .

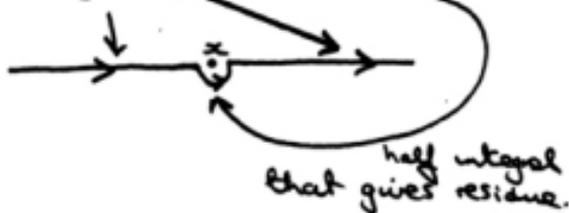
evaluate large semi circle to get 0  
whence

$$f(z) = \frac{1}{2\pi i} \int_{\text{Int}=0} f(t) / (t-z) dt \quad (30.6)$$

Now let  $z$  become real  $= x$  : to be exact as  $1/(t-x) = \infty$  for  $t=x$  we must be more careful. The necessary technique is described in appendix 2 of M. and W. We must write  $z = x + i\epsilon$  to signify that we approach the real axis

from above. ( $\epsilon$  is small  $> 0$ ). Then 132

$$\bullet \frac{1}{t-z} = \frac{1}{t-x-i\epsilon} = P \left\{ \frac{1}{t-x} \right\} + \pi \delta(t-x)$$



$$\text{if } F(x) = \lim_{\substack{z \rightarrow x+i\epsilon \\ \epsilon \rightarrow 0}} F(z)$$

$$\bullet 2\pi i F(x) = P \int_{-\infty}^{\infty} \frac{F(t) dt}{t-x} + \pi F(x)$$

$$\text{or } F(x) = \frac{1}{\pi i} P \int_{-\infty}^{+\infty} \frac{F(t) dt}{t-x}$$

Take real and imaginary parts

$$\text{Re } F(x) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\text{Im } F(t) dt}{t-x} \quad (B10.7)$$

$$\text{Im } F(x) = -\frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\text{Re } F(t) dt}{t-x}$$

• which is as claimed just the Hilbert transform formulae with the

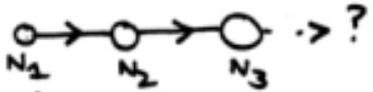
associated function being real and 133  
imaginary parts of the same analytic  
function.

In high energy physics they are important because the basic assumption, causality, is fundamental to all current theories. ~~as~~ they are tested (and hence causality tested) as  $\text{Im} F(t)$  can be (easily) measured as a total cross-section.  $\text{Re} F(x)$  is a harder task and must be measured by coulomb-nuclear interference experiments. Currently we only test the first of Hilbert's relations - so far nature is in accord with dispersion relations.

(iv) Laplace Transform Example

● Lets do an easy example of a Laplace transform. This is M. and W. question 4-14.

"Three radioactive nuclei decay in cascade fashion



$N_i(t)$  are defined by

$$\frac{dN_1}{dt} = -\lambda_1 N_1$$

$$\frac{dN_2}{dt} = \lambda_1 N_1 - \lambda_2 N_2$$

$$\frac{dN_3}{dt} = \lambda_2 N_2 - \lambda_3 N_3$$

Initially  $N_1 = N, N_2 = 0, N_3 = 0.$

Define  $f_i(s) = \int_0^{\infty} N_i(t) e^{-st} dt$

Then  $\int_0^{\infty} \frac{dN_i}{dt} e^{-st} dt = [N_i e^{-st}]_0^{\infty} + s \int_0^{\infty} N_i e^{-st} dt$

$$= s f_i(s) - N_i(0)$$

So applying  $\int_0^{\infty} e^{-st} dt$  to both sides 135

• of the equations we get

$$s f_1 - N = -\lambda_1 f_1$$
$$f_1 = N / (s + \lambda_1)$$

$$s f_2 = \lambda_1 f_1 - \lambda_2 f_2$$

$$\text{or } f_2 = N \lambda_1 / (s + \lambda_1)(s + \lambda_2)$$

$$s f_3 - n = \lambda_2 f_2 - \lambda_3 f_3$$

$$\text{or } (s + \lambda_3) f_3 = n + \lambda_2 f_2$$

$$\text{or } f_3 = \frac{n}{s + \lambda_3} + \frac{N \lambda_1 \lambda_2}{(s + \lambda_1)(s + \lambda_2)}$$

$$N_3(t) = \frac{n}{s + \lambda_3} + \frac{c}{s + \lambda_1} - \frac{c}{s + \lambda_2}$$

$$\text{where } c(\lambda_2 - \lambda_1) = N \lambda_1 \lambda_2$$

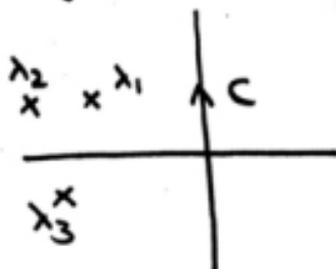
$$\therefore N_3(t) = n e^{-\lambda_3 t} + c [e^{-\lambda_1 t} - e^{-\lambda_2 t}]$$

where we use

136

$$N_3(t) = \frac{1}{2\pi i} \int_C e^{st} f_3(s) ds$$

and we evaluate  
by closing contour  
in the negative Res  
plane (so  $e^{st} \rightarrow 0$ )



and pick up 3 pole residues.