Math 212a Lecture 4

Review.

- **Orthogonal projection.** If M is a complete subspace of a pre- Hilbert space H, then for any $v \in H$ there is a unique $w \in M$ such that $(v-w) \perp M$. This w is characterized as being the unique element of M which minimizes $||v-x||, x \in M$. The idea of the proof is to use the parallelogram law to conclude that if $\{x_n\}$ is a sequence of elements in M for which $||v-x_m||$ approaches the greatest lower bound of $||v-x||$, $x \in M$, then $\{x_m\}$ is a Cauchy sequence. Then the assumption that M is complete guarantees that this sequence has a limit $w \in M$ which minimizes $||v-x||, x \in M$.

The map $v \mapsto w$ is called orthogonal projection of V onto M and will be denoted by π_M.

Projection onto a finite direct sum.

Suppose that the closed subspace M of a pre-Hilbert space is the orthogonal direct sum of a finite number of subspaces

$$M = \bigoplus_i M_i$$

meaning that the M_i are mutually perpendicular and every element x of M can be written as

$$x = \sum x_i, \quad x_i \in M_i.$$

(The orthogonality guarantees that such a decomposition is unique.) Suppose further that each M_i is such that the projection π_{M_i} exists. Then π_M exists and

$$\pi_M(v) = \sum \pi_{M_i}(v). \tag{13}$$

$$\pi_M(v) = \sum \pi_{M_i}(v). \tag{13}$$

Proof. Clearly the right hand side belongs to M. We must show $v - \sum_i \pi_{M_i}(v)$ is orthogonal to every element of M. For this it is enough to show that it is orthogonal to each M_j since every element of M is a sum of elements of the M_j. So suppose $x_j \in M_j$. But $(\pi_{M_i} v, x_j) = 0$ if $i \neq j$. So

$$(v - \sum \pi_{M_i}(v), x_j) = (v - \pi_{M_j}(v), x_j) = 0$$

by the defining property of π_{M_j}.

Review: projection onto a one dimensional subspace.

For example, if M is the one dimensional subspace consisting of all (complex) multiples of a non-zero vector y, then M is complete, since $\mathbf{C}$ is complete. So w exists. Since all elements of M are of the form ay, we can write $w = ay$ for some complex number a. Then $(v - ay, y) = 0$ or

$$(v, y) = a\|y\|^2$$

so

$$a = \frac{(v, y)}{\|y\|^2}.$$

We call a the **Fourier coefficient** of v with respect to y. Particularly useful is the case where $\|y\| = 1$ and we can write

$$a = (v, y). \tag{12}$$

Projection onto a finite dimensional subspace.

$$a = (v, y). \tag{12}$$

$$\pi_M(v) = \sum \pi_{M_i}(v). \tag{13}$$

We now will put the equations (12) and (13) together: Suppose that M is a finite dimensional subspace with an orthonormal basis ϕ_i. This implies that M is an orthogonal direct sum of the one dimensional spaces spanned by the ϕ_i and hence π_M exists and is given by

$$\pi_M(v) = \sum a_i \phi_i \quad \text{where} \quad a_i = (v, \phi_i). \tag{14}$$

Bessel's inequality.

We now look at the infinite dimensional situation and suppose that we are given an orthonormal sequence $\{\phi_i\}_1^\infty$. Any $v \in V$ has its Fourier coefficients

$$a_i = (v, \phi_i)$$

relative to the members of this sequence. Bessel's inequality asserts that

$$\sum_1^\infty |a_i|^2 \leq \|v\|^2, \tag{15}$$

in particular the sum on the left converges.

Proof. Let

$$v_n := \sum_{i=1}^n a_i \phi_i,$$

so that v_n is the projection of v onto the subspace spanned by the first n of the ϕ_i. In any event, $(v - v_n) \perp v_n$ so by the Pythagorean Theorem

$$\|v\|^2 = \|v - v_n\|^2 + \|v_n\|^2 = \|v - v_n\|^2 + \sum_{i=1}^n |a_i|^2.$$

To prove: $$\sum_{1}^{\infty} |a_i|^2 \leq \|v\|^2, \tag{15}$$

Let
$$v_n := \sum_{i=1}^{n} a_i \phi_i,$$

so that v_n is the projection of v onto the subspace spanned by the first n of the ϕ_i. In any event, $(v - v_n) \perp v_n$ so by the Pythagorean Theorem

$$\|v\|^2 = \|v - v_n\|^2 + \|v_n\|^2 = \|v - v_n\|^2 + \sum_{i=1}^{n} |a_i|^2.$$

This implies that
$$\sum_{i=1}^{n} |a_i|^2 \leq \|v\|^2$$

and letting $n \to \infty$ shows that the series on the left of Bessel's inequality converges and that Bessel's inequality holds.

Parseval's equation.

Continuing the above argument, observe that

$$\|v - v_n\|^2 \to 0 \;\Leftrightarrow\; \sum |a_i|^2 = \|v\|^2.$$

But $\|v - v_n\|^2 \to 0$ if and only if $\|v - v_n\| \to 0$ which is the same as saying that $v_n \to v$. But v_n is the n-th partial sum of the series $\sum a_i\phi_i$, and in the language of series, we say that a series converges to a limit v and write $\sum a_i\phi_i = v$ if and only if the partial sums approach v. So

$$\sum a_i\phi_i = v \quad \Leftrightarrow \quad \sum_i |a_i|^2 = \|v\|^2. \tag{16}$$

In general, we will call the series $\sum_i a_i\phi_i$ the Fourier series of v (relative to the given orthonormal sequence) whether or not it converges to v. Thus Parseval's equality says that the Fourier series of v converges to v if and only if $\sum |a_i|^2 = \|v\|^2$.

Orthonormal bases.

We still suppose that V is merely a pre-Hilbert space. We say that an orthonormal sequence $\{\phi_i\}$ is a **basis** of V if every element of V is the sum of its Fourier series. For example, one of our tasks will be to show that the exponentials $\{e^{inx}\}_{n=-\infty}^{\infty}$ form a basis of $\mathcal{C}(\mathbf{T})$.

If the orthonormal sequence ϕ_i is a basis, then any v can be approximated as closely as we like by finite linear combinations of the ϕ_i, in fact by the partial sums of its Fourier series. We say that the finite linear combinations of the ϕ_i are *dense* in V. Conversely, suppose that the finite linear combinations of the ϕ_i are dense in V. This means that for any v and any $\epsilon > 0$ we can find an n and a set of n complex numbers b_i such that

$$\|v - \sum b_i \phi_i\| \leq \epsilon.$$

But we know that v_n is the closest vector to v among all the linear combinations of the first n of the ϕ_i. so we must have

$$\|v - v_n\| \leq \epsilon.$$

$$\|v - v_n\| \leq \epsilon.$$

But this says that the Fourier series of v converges to v, i.e. that the ϕ_i form a basis. For example, we know from Fejer's theorem that the exponentials e^{ikx} are dense in $\mathcal{C}(\mathbf{T})$. Hence we know that they form a basis of the pre-Hilbert space $\mathcal{C}(\mathbf{T})$. We will give some alternative proofs of this fact below.

Orthonormal bases, 2.

In the case that V is actually a Hilbert space, and not merely a pre-Hilbert space, there is an alternative and very useful criterion for an orthonormal sequence to be a basis: Let M be the set of all limits of finite linear combinations of the ϕ_i. Any Cauchy sequence in M converges (in V) since V is a Hilbert space, and this limit belongs to M since it is itself a limit of finite linear combinations of the ϕ_i (by the diagonal argument for example). Thus $V = M \oplus M^\perp$, and the ϕ_i form a basis of M. So the ϕ_i form a basis of V if and only if $M^\perp = \{0\}$. But this is the same as saying that no non-zero vector is orthogonal to all the ϕ_i. So we have proved

Proposition 1 *In a Hilbert space, the orthonormal set* $\{\phi_i\}$ *is a basis if and only if no non-zero vector is orthogonal to all the* ϕ_i.

Symmetric transformations.

We continue to let V denote a pre-Hilbert space. Let T be a linear transformation of V into itself. This means that for every $v \in V$ the vector $Tv \in V$ is defined and that Tv depends linearly on v : $T(av + bw) = aTv + bTw$ for any two vectors v and w and any two complex numbers a and b. We recall from linear algebra that a non-zero vector v is called an eigenvector of T if Tv is a scalar times v, in other words if $Tv = \lambda v$ where the number λ is called the corresponding eigenvalue.

A linear transformation T on V is called **symmetric** if for any pair of elements v and w of V we have

$$(Tv, w) = (v, Tw).$$

Notice that if v is an eigenvector of a symmetric transformation T with eigenvalue λ, then

$$\lambda(v, v) = (\lambda v, v) = (Tv, v) = (v, Tv) = (v, \lambda v) = \overline{\lambda}(v, v),$$

so $\lambda = \overline{\lambda}$. In other words, all eigenvalues of a symmetric transformation are real.

Bounded linear transformations.

We will let $\mathbf{S} = \mathbf{S}(V)$ denote the "unit sphere" of V, i.e. $\mathbf{S}$ denotes the set of all $\phi \in V$ such that $\|\phi\| = 1$. A linear transformation T is called **bounded** if $\|T\phi\|$ is bounded as ϕ ranges over all of $\mathbf{S}$. If T is bounded, we let

$$\|T\| := \max_{\phi \in \mathbf{S}} \|T\phi\|.$$

Then

$$\|Tv\| \leq \|T\|\|v\|$$

for all $v \in V$. A linear transformation on a finite dimensional space is automatically bounded, but not so for an infinite dimensional space.

The null space (kernel) and the range.

Also, for any linear transformation T, we will let $N(T)$ denote the kernel of T, so

$$N(T) = \{v \in V | Tv = 0\}$$

and $R(T)$ denote the range of T, so

$$R(T) := \{v | v = Tw \text{ for some } w \in V\}.$$

Both $N(T)$ and $R(T)$ are linear subspaces of V.

“self-adjoint” vs.”symmetric.

For bounded transformations, the phrase “self-adjoint” is synonymous with “symmetric”. Later on we will need to study non-bounded (not everywhere defined) symmetric transformations, and then a rather subtle and important distinction will be made between self-adjoint transformations and those which are merely symmetric. But for the rest of this section we will only be considering bounded linear transformations, and so we will freely use the phrase “self-adjoint”, and (usually) drop the adjective “bounded” since all our transformations will be assumed to be bounded.

We denote the set of all (bounded) self-adjoint transformations by $\mathcal{A}$, or by $\mathcal{A}(V)$ if we need to make V explicit.

Non-negative self-adjoint transformations.

If T is a self-adjoint transformation, then

$$(Tv, v) = (v, Tv) = \overline{(Tv, v)}$$

so (Tv, v) is always a real number. More generally, for any pair of elements v and w,

$$(Tv, w) = \overline{(Tw, v)}.$$

Since (Tv, w) depends linearly on v for fixed w, we see that the rule which assigns to every pair of elements v and w the number (Tv, w) satisfies the first two conditions in our definition of a semi-scalar product. Since (Tv, v) might be negative, condition 3. of the definition need not be satisfied. This leads to the following definition:

A self-adjoint transformation T is called **non-negative** if

$$(Tv, v) \geq 0 \quad \forall\ v \in V.$$

Bounded non-negative self-adjoint operators.

So if T is a non-negative self-adjoint transformation, then the rule which assigns to every pair of elements v and w the number (Tv, w) *is* a semi-scalar product to which we may apply the Cauchy-Schwarz inequality and conclude that

$$|(Tv, w)| \leq (Tv, v)^{\frac{1}{2}} (Tw, w)^{\frac{1}{2}}.$$

Now let us assume in addition that T is bounded with norm $\|T\|$. Let us take $w = Tv$ in the preceding inequality. We get

$$\|Tv\|^2 = |(Tv, Tv)| \leq (Tv, v)^{\frac{1}{2}} (TTv, Tv)^{\frac{1}{2}}.$$

Now apply the Cauchy-Schwarz inequality for the original scalar product to the last factor on the right:

$$(TTv, Tv)^{\frac{1}{2}} \leq \|TTv\|^{\frac{1}{2}} \|Tv\|^{\frac{1}{2}} \leq \|T\|^{\frac{1}{2}} \|Tv\|^{\frac{1}{2}} \|Tv\|^{\frac{1}{2}} = \|T\|^{\frac{1}{2}} \|Tv\|,$$

where we have used the defining property of $\|T\|$ in the form $\|TTv\| \leq \|T\|\|Tv\|$. Substituting this into the previous inequality we get

$$\|Tv\|^2 \leq (Tv, v)^{\frac{1}{2}} \|T\| \|Tv\|.$$

$$\|Tv\|^2 \leq (Tv,v)^{\frac{1}{2}}\|T\|\|Tv\|.$$

If $\|Tv\| \neq 0$ we may divide this inequality by $\|Tv\|$ to obtain

$$\|Tv\| \leq \|T\|^{\frac{1}{2}}(Tv,v)^{\frac{1}{2}}. \tag{17}$$

This inequality is clearly true if $\|Tv\| = 0$ and so holds in all cases.

We will make much use of this inequality. For example, it follows from (17) that

$$(Tv,v) = 0 \;\; \Rightarrow \;\; Tv = 0. \tag{18}$$

It also follows from (17) that if we have a sequence $\{v_n\}$ of vectors with $(Tv_n, v_n) \to 0$ then $\|Tv_v\| \to 0$ and so

$$(Tv_n, v_n) \to 0 \;\; \Rightarrow \;\; Tv_n \to 0. \tag{19}$$

Notice that if T is a bounded self adjoint transformation, not necessarily non-negative, then $rI - T$ is a non-negative self-adjoint transformation if $r \geq \|T\|$: Indeed,

$$((rI-T)v,v) = r(v,v) - (Tv,v) \geq (r - \|T\|)(v,v) \geq 0$$

since, by Cauchy-Schwarz,

$$(Tv,v) \leq |(Tv,v)| \leq \|Tv\|\|v\| \leq \|T\|\|v\|^2 = \|T\|(v,v).$$

So we may apply the preceding results to $rI - T$.

Compact self-adjoint transformations.

We say that the self-adjoint transformation T is **compact** if it has the following property: Given any sequence of elements $u_n \in \mathbf{S}$, we can choose a subsequence u_{n_i} such that the sequence Tu_{n_i} converges to a limit in V.

Some remarks about this complicated looking definition: In case V is finite dimensional, every linear transformation is bounded, hence the sequence Tu_n lies in a bounded region of our finite dimensional space, and hence by the completeness property of the real (and hence complex) numbers, we can always find such a convergent subsequence. So in finite dimensions every T is compact. More generally, the same argument shows that if $R(T)$ is finite dimensional and T is bounded then T is compact. So the definition is of interest essentially in the case when $R(T)$ is infinite dimensional.

Also notice that if T is compact, then T is bounded. Otherwise we could find a sequence u_n of elements of $\mathbf{S}$ such that $\|Tu_n\| \geq n$ and so no subsequence Tu_{n_i} can converge.

Eigenvalues of compact self-adjoint operators.

We now come to the key result which we will use over and over again:

Theorem 3 *Let T be a compact self-adjoint operator. Then $R(T)$ has an orthonormal basis $\{\phi_i\}$ consisting of eigenvectors of T and if $R(T)$ is infinite dimensional then the corresponding sequence $\{r_n\}$ of eigenvalues converges to* 0.

Proof. We know that T is bounded. If $T = 0$ there is nothing to prove. So assume that $T \neq 0$ and let

$$m_1 := \|T\| > 0.$$

By the definition of $\|T\|$ we can find a sequence of vectors $u_n \in \mathbf{S}$ such that $\|Tu_n\| \to \|T\|$. By the definition of compactness we can find a subsequence of this sequence so that $Tu_{n_i} \to w$ for some $w \in V$. On the other hand, the transformation T^2 is self-adjoint and bounded by $\|T\|^2$. Hence $\|T\|^2 I - T^2$ is non-negative, and

$$((\|T\|^2 I - T^2)u_n, u_n) = \|T\|^2 - \|Tu_n\|^2 \to 0.$$

So we know from (19) that

$$\|T\|^2 u_n - T^2 u_n \to 0.$$

Passing to the subsequence we have $T^2 u_{n_i} = T(Tu_{n_i}) \to Tw$ and so

$$\|T\|^2 u_{n_i} \to Tw$$

or

$$u_{n_i} \to \frac{1}{m_1^2} Tw.$$

$$u_{n_i} \to \frac{1}{m_1^2} T w.$$

Applying T to this we get

$$T u_{n_i} \to \frac{1}{m_1^2} T^2 w$$

or

$$T^2 w = m_1^2 w.$$

Also $\|w\| = \|T\| = m_1 \neq 0$. So $w \neq 0$. So w is an eigenvector of T^2 with eigenvalue m_1^2. We have

$$0 = (T^2 - m_1^2) w = (T + m_1)(T - m_1) w.$$

$$0 = (T^2 - m_1^2)w = (T + m_1)(T - m_1)w.$$

If $(T - m_1)w = 0$, then w is an eigenvector of T with eigenvalue m_1 and we normalize by setting

$$\phi_1 := \frac{1}{\|w\|}w.$$

Then $\|\phi_1\| = 1$ and

$$T\phi_1 = m_1\phi_1.$$

If $(T - m_1)w \neq 0$ then $y := (T - m_1)w$ is an eigenvector of T with eigenvalue $-m_1$ and again we normalize by setting

$$\phi_1 := \frac{1}{\|y\|}y.$$

So we have found a unit vector $\phi_1 \in R(T)$ which is an eigenvector of T with eigenvalue $r_1 = \pm m_1$.

Now let

$$V_2 := \phi_1^\perp.$$

If $(w, \phi_1) = 0$, then

$$(Tw, \phi_1) = (w, T\phi_1) = r_1(w, \phi_1) = 0.$$

In other words,

$$T(V_2) \subset V_2$$

and we can consider the linear transformation T restricted to V_2 which is again compact. If we let m_2 denote the norm of the linear transformation T when restricted to V_2 then $m_2 \leq m_1$ and we can apply the preceding procedure to find a unit eigenvector ϕ_2 with eigenvalue $\pm m_2$.

We proceed inductively, letting

$$V_n := \{\phi_1, \dots, \phi_{n-1}\}^\perp$$

and find an eigenvector ϕ_n of T restricted to V_n with eigenvalue $\pm m_n \neq 0$ if the restriction of T to V_n is not zero. So there are two alternatives:

- after some finite stage the restriction of T to V_n is zero. In this case $R(T)$ is finite dimensional with orthonormal basis $\phi_1, \dots, \phi_{n-1}$. Or
- The process continues indefinitely so that at each stage the restriction of T to V_n is not zero and we get an infinite sequence of eigenvectors and eigenvalues r_i with $|r_i| \geq |r_{i+1}|$.

The first case is one of the alternatives in the theorem, so we need to look at the second alternative.

We first prove that $|r_n| \to 0$. If not, there is some $c > 0$ such that $|r_n| \geq c$ for all n (since the $|r_n|$ are decreasing). If $i \neq j$,then by the Pythagorean theorem we have

$$\|T\phi_i - T\phi_j\|^2 = \|r_i\phi_i - r_j\phi_j\|^2 = r_i^2\|\phi_i\|^2 + r_j^2\|\phi_j\|^2.$$

Since $\|\phi_i\| = \|\phi_j| = 1$ this gives

$$\|T\phi_i - T\phi_j\|^2 = r_i^2 + r_j^2 \geq 2c^2.$$

Hence no subsequence of the $T\phi_i$ can converge, since all these vectors are at least a distance $c\sqrt{2}$ apart. This contradicts the compactness of T.

To complete the proof of the theorem we must show that the ϕ_i form a basis of $R(T)$. So if $w = Tv$ we must show that the Fourier series of w with respect to the ϕ_i converges to w. We begin with the Fourier coefficients of v relative to the ϕ_i which are given by

$$a_n = (v, \phi_n).$$

Then the Fourier coefficients of w are given by

$$b_i = (w, \phi_i) = (Tv, \phi_i) = (v, T\phi_i) = (v, r_i\phi_i) = r_i a_i.$$

So

$$w - \sum_{i=1}^{n} b_i\phi_i = Tv - \sum_{i=1}^{n} a_i r_i \phi_i = T(v - \sum_{i=1}^{n} a_i\phi_i).$$

Now $v - \sum_{i=1}^{n} a_i\phi_i$ is orthogonal to $\phi_1, \ldots, \phi_n$ and hence belongs to V_{n+1}. So

$$\|T(v - \sum_{i=1}^{n} a_i\phi_i)\| \leq |r_{n+1}| \|(v - \sum_{i=1}^{n} a_i\phi_i)\|.$$

Completion of the proof.

$$w - \sum_{i=1}^{n} b_i \phi_i = Tv - \sum_{i=1}^{n} a_i r_i \phi_i = T(v - \sum_{i=1}^{n} a_i \phi_i).$$

$$\|T(v - \sum_{i=1}^{n} a_i \phi_i)\| \leq |r_{n+1}| \|(v - \sum_{i=1}^{n} a_i \phi_i)\|.$$

By the Pythagorean theorem,

$$\|(v - \sum_{i=1}^{n} a_i \phi_i)\| \leq \|v\|.$$

Putting the two previous inequalities together we get

$$\|w - \sum_{i=1}^{n} b_i \phi_i\| = \|T(v - \sum_{i=1}^{n} a_i \phi_i)\| \leq |r_{n+1}| \|v\| \to 0.$$

This proves that the Fourier series of w converges to w concluding the proof of the theorem.

A converse to the key theorem.

The "converse" of the above result is easy. Here is a version: Suppose that $\mathbf{H}$ is a Hilbert space with an orthonormal basis $\{\phi_i\}$ consisting of eigenvectors of an operator T, so $T\phi_i = \lambda_i \phi_i$, and suppose that $\lambda_i \to 0$ as $i \to \infty$. Then T is compact. Indeed, for each j we can find an $N = N(j)$ such that

$$|\lambda_r| < \frac{1}{j} \ \ \forall \ r > N(j).$$

We can then let $\mathbf{H}_j$ denote the closed subspace spanned by all the eigenvectors $\phi_r, r > N(j)$, so that

$$\mathbf{H} = \mathbf{H}_j^{\perp} \oplus \mathbf{H}_j$$

is an orthogonal decomposition and $\mathbf{H}_j^{\perp}$ is finite dimensional, in fact is spanned the first $N(j)$ eigenvectors of T.

Now let $\{u_i\}$ be a sequence of vectors with $\|u_i\| \leq 1$ say. We decompose each element as

$$u_i = u_i' \oplus u_i'', \qquad u_i' \in \mathbf{H}_1^{\perp},\ u_i'' \in \mathbf{H}_j.$$

We can choose a subsequence so that u_{i_k}' converges, because they all belong to a finite dimensional space, and hence so does Tu_{i_k} since T is bounded. We can decompose every element of this subsequence into its $\mathbf{H}_2^{\perp}$ and $\mathbf{H}_2$ components, and choose a subsequence so that the first component converges. Proceeding in this way, and then using the Cantor diagonal trick of choosing the k-th term of the k-th selected subsequence, we have found a subsequence such that for any fixed j, the (now relabeled) subsequence, the $\mathbf{H}_j^{\perp}$ component of Tu_j converges. But the $\mathbf{H}_j$ component of Tu_j has norm less than $1/j$, and so the sequence converges by the triangle inequality.

Fourier's Fourier series.

We want to apply the theorem about compact self-adjoint operators that we proved in the preceding section to conclude that the functions e^{inx} form an orthonormal basis of the space $\mathcal{C}(\mathbf{T})$. In fact, a direct proof of this fact is elementary, using integration by parts. So we will

pause to given this direct proof. Then we will go back and give a (more complicated) proof of the same fact using our theorem on compact operators. The reason for giving the more complicated proof is that it extends to far more general situations.

We have let $\mathcal{C}(\mathbf{T})$ denote the space of continuous functions on the real line which are periodic with period 2π. We will let $\mathcal{C}^1(\mathbf{T})$ denote the space of periodic functions which have a continuous first derivative (necessarily periodic) and by $\mathcal{C}^2(\mathbf{T})$ the space of periodic functions with two continuous derivatives. If f and g both belong to $\mathcal{C}^1(\mathbf{T})$ then integration by parts gives

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f'\overline{g}dx = -\frac{1}{2\pi}\int_{-\pi}^{\pi} f\overline{g}'dx$$

since the boundary terms, which normally arise in the integration by parts formula, cancel, due to the periodicity of f and g. If we take $g = e^{inx}/(in), n \neq 0$ the integral on the right hand side of this equation is the Fourier coefficient:

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f'\overline{g}dx = -\frac{1}{2\pi}\int_{-\pi}^{\pi} f\overline{g}'dx$$

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

We thus obtain

$$c_n = \frac{1}{in}\frac{1}{2\pi}\int_{-\pi}^{\pi} f'(x)e^{-inx}dx$$

so, for $n \neq 0$,

$$|c_n| \leq \frac{A}{n} \quad \text{where } A := \frac{1}{2\pi}\int_{-\pi}^{\pi} |f'(x)|dx$$

is a constant independent of n (but depending on f).

If $f \in \mathcal{C}^2(\mathbf{T})$ we can take $g(x) = -e^{inx}/n^2$ and integrate by parts twice. We conclude that (for $n \neq 0$)

$$|c_n| \leq \frac{B}{n^2} \quad \text{where} \quad B := \frac{1}{2\pi}\int_{-\pi}^{\pi} |f''(x)|^2$$

If $f \in \mathcal{C}^2(\mathbf{T})$ we can take $g(x) = -e^{inx}/n^2$ and integrate by parts twice. We conclude that (for $n \neq 0$)

$$|c_n| \leq \frac{B}{n^2} \quad \text{where} \quad B := \frac{1}{2\pi} \int_{-\pi}^{\pi} |f''(x)|^2$$

is again independent of n. But this proves that the Fourier series of f,

$$\sum c_n e^{inx}$$

converges uniformly and absolutely for and $f \in \mathcal{C}^2(\mathbf{T})$. The limit of this series is therefore some continuous periodic function. We must prove that this limit equals f. So we must prove that at each point f

$$\sum c_n e^{iny} \to f(y).$$

Replacing $f(x)$ by $f(x-y)$ it is enough to prove this formula for the case $y=0$. So we must prove that for any $f \in \mathcal{C}^2(\mathbf{T})$ we have

$$\lim_{N,M\to\infty} \sum_{-N}^{M} c_n \to f(0).$$

Write $f(x) = (f(x) - f(0)) + f(0)$. The Fourier coefficients of any constant function c all vanish except for the c_0 term which equals c. So the above limit is trivially true when f is a constant. Hence, in proving the above formula, it is enough to prove it under the additional assumption that $f(0) = 0$, and we need to prove that in this case

$$\lim_{N,M\to\infty} (c_{-N} + c_{-N+1} + \cdots + c_M) \to 0.$$

So the above limit is trivially true when f is a constant. Hence, in proving the above formula, it is enough to prove it under the additional assumption that $f(0) = 0$, and we need to prove that in this case

$$\lim_{N,M\to\infty} (c_{-N} + c_{-N+1} + \cdots + c_M) \to 0.$$

The expression in parenthesis is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g_{N,M}(x)}dx$$

where

$$g_{N,M}(x) = e^{-iNx}+e^{-i(N-1)x}+\cdots+e^{iMx} = e^{-iNx}\left(1 + e^{ix} + \cdots + e^{i(M+N)x}\right) =$$

$$e^{-iNx}\frac{1 - e^{i(M+N+1)x}}{1 - e^{ix}} = \frac{e^{-iNx} - e^{i(M+1)x}}{1 - e^{ix}}, \quad x \neq 0$$

$$(c_{-N} + c_{-N+1} + \cdots + c_M)$$

The expression in parenthesis is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g_{N,M}(x)}dx = \frac{e^{-iNx} - e^{i(M+1)x}}{1 - e^{ix}}, \quad x \neq 0$$

where we have used the formula for a geometric sum. By l'Hôpital's rule, this extends continuously to the value $M + N + 1$ for $x = 0$. Now $f(0) = 0$, and since f has two continuous derivatives, the function

$$h(x) := \frac{f(x)}{1 - e^{-ix}}$$

defined for $x \neq 0$ (or any multiple of 2π) extends, by l'Hôpital's rule, to a function defined at all values, and which is continuously differentiable and periodic. Hence the limit we are computing is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} h(x)e^{iNx}dx - \frac{1}{2\pi}\int_{-\pi}^{\pi} h(x)e^{-i(M+1)x}dx$$

and we know that each of these terms tends to zero.

We have thus proved that the Fourier series of any twice differentiable periodic function converges uniformly and absolutely to that function. If we consider the space $\mathcal{C}^2(\mathbf{T})$ with our usual scalar product

$$(f,g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f\overline{g}dx$$

then the functions e^{inx} are dense in this space, since uniform convergence implies convergence in the $\| \ \|$ norm associated to $(\ , \ )$. So, on general principles, Bessel's inequality and Parseval's equation hold.

It is not true in general that the Fourier series of a continuous function converges uniformly to that function (or converges at all in the sense of uniform convergence). However it is true that we *do* have convergence in the L_2 norm, i.e. the Hilbert space $\| \ \|$ norm on $\mathcal{C}(\mathbf{T})$. To prove this, we need only prove that the exponential functions e^{inx} are dense, and since they are dense in $\mathcal{C}^2(\mathbf{T})$, it is enough to prove that $\mathcal{C}^2(\mathbf{T})$ is dense in $\mathcal{C}(\mathbf{T})$. For this, let ϕ be a function defined on the line with at least two continuous bounded derivatives with $\phi(0) = 1$ and of total integral equal to one and which vanishes rapidly at infinity.

A favorite is the Gauss normal function

$$\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Equally well, we could take ϕ to be a function which actually vanishes outside of some neighborhood of the origin. Let

$$\phi_t(x) := \frac{1}{t} \phi\left(\frac{x}{t}\right).$$

As $t \to 0$ the function ϕ_t becomes more and more concentrated about the origin, but still has total integral one. Hence, for any bounded continuous function f, the function $\phi_t \star f$ defined by

$$(\phi_t \star f)(x) := \int_{-\infty}^{\infty} f(x-y)\phi_t(y)dy = \int_{-\infty}^{\infty} f(u)\phi_t(x-u)du.$$

satisfies $\phi_t \star f \to f$ uniformly on any finite interval.

$$(\phi_t \star f)(x) := \int_{-\infty}^{\infty} f(x-y)\phi_t(y)dy = \int_{-\infty}^{\infty} f(u)\phi_t(x-u)du.$$

satisfies $\phi_t \star f \longrightarrow f$ uniformly on any finite interval. From the rightmost expression for $\phi_t \star f$ above we see that $\phi_t \star f$ has two continuous derivatives. From the first expression we see that $\phi_t \star f$ is periodic if f is. This proves that $\mathcal{C}^2(\mathbf{T})$ is dense in $\mathcal{C}(\mathbf{T})$. We have thus proved convergence in the L_2 norm.