Plancherel:

We have $X = C[0,1]$, complex valued continuous functions on $[0,1]$ with metric
$$d(f,g) = \sqrt{\int_0^1 |f-g|^2}$$
We have l^2, space of (two sided) infinite sequences which are square summable with metric
$$d(a,b) = \sqrt{\sum |a_k - b_k|^2}.$$
We have map from X to Y
$$Tf = \{\hat{f}_k : -\infty < n < \infty\}.$$
We have the set D of all functions that satisfy the convergence theorem. That is, all C^1 functions on $[0,1]$ with $f(0) = f(1)$ and $f'(0) = f'(1)$.

We claim that D is dense in $C[0,1]$. Remember the metric is d above, not the sup metric. Of course, D can not be dense in sup metric. This is so because convergence in sup metric, in particular, implies point-wise convergence and so period one stays in the limit (though not differentiability).

Let f be given. Assume that f is real valued. Let $\epsilon > 0$ be given. By Weierstrass we can get a polynomial P with sup distance from f smaller than ϵ. Since length of interval of integration is one, we see $d(f,P) < \epsilon$.

To continue, here is an observation. Given $-\infty < a < b < c < d < \infty$ there is a C^1 function φ on R such that $\varphi(x)$ is
zero for $x \leq a$;
then increases to one as x increases to b;
then remains one up to c;
then decreases to zero as x further increases to d;
then remains zero for $x \geq d$.
Such a function, actually C^∞ function was constructed last year. But since we do not need C^∞ now, we can do it in a simpler way as follows.

Define ψ on R:

zero below a;

between a and b its graph is an isosceles triangle with base $[a, b]$ and height $+2/(b-a)$ thus graph is above x-axis;

again zero between b and c;

between c and d its graph is an isosceles triangle with base $[c, d]$ and height $-2/(d-c)$ thus graph is below x-axis;

from d onwards zero again.

Now take indefinite integral

$$\varphi(x) = \int_{-\infty}^{x} \psi(t)dt.$$

By fundamental theorem of calculus, $\varphi' = \psi$ and is hence continuous and φ has all the required properties.

Now to continue with our earlier argument, let P be a real polynomial on $[0, 1]$ Suppose $|P| \leq M$. Take φ of the above paragraph with

$$a = 0; \quad b = 1/8M; \quad c = 1 - (1/8M); \quad d = 1$$

and take the product $g = P\varphi$. Observe that g and g' are continuous and periodic; g and P agree on a large part of the interval (namely, on $[b, c]$) and direct calculation shows you $d(P, g) < \epsilon$

By triangle inequality, we have proved the following. Given any real continuous f on $[0, 1]$, there is $g \in D$ with $d(f, g) < \epsilon$. Complex case follows. If $f = f_1 + if_2$ complex valued, get g_1 and g_2 for f_1 and f_2 with $\epsilon/2$ and argue that by taking $g = g_1 + ig_2$ we have $d(f, g) < \epsilon$ and $g \in D$.

All this goes to show that D is a dense subset of $C[0, 1]$.

Observe that we have already shown, after Bessel inequality,

$$\sum |\hat{f}_k|^2 \leq \int |f|^2$$

Since $T(f - g) = Tf - Tg$, conclude

$$d(Tf, Tg) \leq d(f, g)$$

In other words, T is a continuous map. But remember that T is an isometry on D. If you now take $f, g \in C[0, 1]$; you can get $f_n \in D$; $f_n \to f$ and $g_n \in D$; $g_n \to g$. Since both T and the distance function are continuous, we get

$$d(Tf, Tg) = \lim d(Tf_n, Tg_n) = \lim d(f_n, g_n) = d(f, g)$$

This shows that T is an isometry on all of X into Y.

since $T0 = 0$, taking $g = 0$ we get Plancherel for all $f \in C[0,1]$.

Thus we have completed proof of Plancherel identity and thereby completed proof of the fact $\sum(1/k^2) = \pi^2/6$.

Jacobi identity:

For real number $t > 0$ define

$$\vartheta(t) = \sum_{-\infty}^{\infty} e^{-n^2\pi t.}$$

This is called theta function. This series is convergent because $\exp\{-\pi t\} < 1$ and geometric series $\sum a^n$ converges and the above series is dominated by this geometric series.

This function appears in many contexts: Riemann zeta function, number theory, statistical physics. Here is an useful identity due to Jacobi.

$$\vartheta(t) = \frac{1}{\sqrt{t}}\,\vartheta(1/t).$$

We shall prove this identity now. Consider the function

$$f(x) = \sum_{-\infty}^{\infty} e^{-(x-n)^2/2t}; \qquad 0 \le x \le 1.$$

This series is uniformly convergent. In fact, fix any $x \in [0,1]$.
If $|n| = 2$, then $(x-n)^2 \ge 1^2$; if $|n| = 3$, then $(x-n)^2 \ge 2^2$; and so on; thus this series is dominated (beyond the terms $n = 0, \pm 1$) again by a convergent series of numbers and is hence uniformly convergent.

In particular this series defines a continuous function on $[0,1]$. This is also periodic, $f(0) = f(1)$; it is the same sum, just the terms get shifted. Let us compute its Fourier coefficients.

$$\hat{f}_k = \int_0^1 \sum_{-\infty}^{\infty} e^{-(x-n)^2/2t} e^{-2\pi i k x} dx;$$

Because of uniform convergence you can interchange the order of integration and summation;

$$= \sum_{-\infty}^{\infty} \int_0^1 e^{-(x-n)^2/2t} e^{-2\pi i k x} dx;$$

Now a change of variable and the realization that $\exp\{2\pi ikn\} = 1$ for any integer n will tell us the following. For example take the term $n = -1$

$$\int_0^1 e^{-(x+1)^2/2t} e^{-2\pi ikx} dx = \int_1^2 e^{-y^2/2t} e^{-2\pi iky} dy$$

(too lazy to write for general $\pm n$) we see

$$\hat{f}_k = \int_{-\infty}^{\infty} e^{-y^2/2t} e^{-2\pi iky} dy.$$

Fortunately, this is something we had already known, done under the name characteristic function of the normal distribution. Recall

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} e^{iux} dx = e^{-u^2/2}.$$

Changing the variable x to $y/\sqrt{t}$ we get

$$\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-y^2/2t} e^{iuy/\sqrt{t}} dy = e^{-u^2/2}.$$

Remember $t > 0$ is fixed. Use this formula with $u = -2\pi k\sqrt{t}$ to see

$$\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-y^2/2t} e^{-2\pi ikty} dy = e^{-4\pi^2 k^2 t/2} = e^{-2\pi^2 k^2 t}.$$

or

$$\int_{-\infty}^{\infty} e^{-y^2/2t} e^{-2\pi ikty} dy = \sqrt{2\pi t} \;\; e^{-2\pi^2 k^2 t}.$$

Thus returning to our calculation

$$\hat{f}_k = \sqrt{2\pi t} \;\; e^{-2\pi^2 k^2 t}.$$

What an achievement! we have been able to evaluate a rather complicated looking integral, the integrand is itself an infinite series.

Suppose that someone tells us that the Fourier expansion is valid for our function, that is

$$f(x) = \sum_{-\infty}^{\infty} \hat{f}_n \;\; e^{2\pi inx}.$$

Then we see that for every $x \in [0, 1]$

$$\sum_{-\infty}^{\infty} e^{-(x-n)^2/2t} = \sum_{-\infty}^{\infty} \sqrt{2\pi t} \;\; e^{-2\pi^2 n^2 t} e^{2\pi inx}.$$

Read this equation for $x = 0$ to get

$$\sum_{-\infty}^{\infty} e^{-n^2/2t} = \sum_{-\infty}^{\infty} \sqrt{2\pi t} e^{-2\pi^2 n^2 t}.$$

or

$$\frac{1}{\sqrt{2\pi t}} \sum_{-\infty}^{\infty} e^{-n^2/2t} = \sum_{-\infty}^{\infty} e^{-2\pi^2 n^2 t}.$$

read this equation for $t/2\pi$ instead of t.

$$\frac{1}{\sqrt{t}} \sum_{-\infty}^{\infty} e^{-n^2\pi/t} = \sum_{-\infty}^{\infty} e^{-n^2\pi t}.$$

This is precisely what we are looking for.

The validity of Fourier expansion can be justified by differentiating the series for $f(x)$ term by term and showing that the series so obtained is uniformly convergent. Recall that if the term-by-term derived series is uniformly convergent then the original series can be differentiated term-by-term. We also can see that the derived series is continuous and period one function. All this just depends on the fact that when $0 < a < 1$, the series $\sum_{n\geq 1} na^n$ converges. However we shall not pause to verify the details.

Returning to the first convergence theorem, it is also possible to prove convergence theorem for not necessarily smooth functions. Let f be a continuous periodic function. Define the partial sums S_N as earlier for $N = 0, 1, 2, \cdots$. Now take their averages

$$\sigma_n = \frac{1}{n}\{S_0 + S_1 + \cdots + s_{n-1}\}$$

Then σ_n converges uniformly to f on $[0, 1]$. This is known as Fejer's theorem and proved by realizing that

$$\sigma_n(x) = \int f(y) F_n(x-y) dy$$

where F_n is Fejer kernel

$$F_n(\theta) = \frac{1}{n}\left[\frac{\sin n\pi\theta}{\sin \pi\theta}\right]^2$$

As you may have noticed we are dealing with expressions like

$$\int f(y) K(x-y) dy.$$

This leads to a general theory called convolutions. Thus the above function of x is called convolution of the two functions f and K.

Thus S_N is convolution of f and D_N; whereas σ_n is convolution of f and F_n. The point is that these functions D_N or F_n put more and more 'mass' at the point zero as N or n becomes large; thus when you translate by x they put more and more mass at x and thus capture the value of f at x. This concept has a beautiful theory behind it that goes by the name of approximate identities. But we shall discuss no more.

This completes our discussion of Fourier series.

HA:

Three problems from home assignments need to be sorted out.

1. ($\spadesuit$) :

There is again an error, now in the last home work problem.

As pointed out by Uma, it is not a metric. Here is the idea, if you have two points on the circle, their distance is the length of the 'smaller part' of the arc between them. My error lies in saying that for points in $(0,1)$ the distance is the good old one. No.

Here is the correction and very brief solution.

Let $X = [0,1]$ usual metric. Let $Y = (0,1) \cup \{\spadesuit\}$ Thus Y has all points of X except zero and one, instead it has one extra point. This is the bag containing zero and one. Here is the metric

$$d^*(s,t) = \min\{|t-s|;\ \ s \wedge t + 1 - (s \vee t)\}; \quad 0 < s,t < 1$$

$$d^*(s,\spadesuit) = d^*(\spadesuit, s) = \min\{|s-0|, |s-1|\} = s \wedge (1-s); \quad 0 < s < 1.$$

$$d^*(\spadesuit,\spadesuit) = 0.$$

If $0 < s < t < 1$ then you can go from s to t directly travelling distance $t-s$ or you can go from s to 0 which is same as 1 and then to t, you get $s + (1-t)$. This explains the first formula. The second formula is similarly explained: to go from s to 0 which is same as 1, take the shortest route.

Note that in the definition of d^* the two quantities whose minimum is being taken add to one so that the distance is always at most $1/2$.

The questions are: Does this satisfy rules for being distance function? Is the space homeomorphic to the circle? There appear to be several smart ways.

Probably, the best way is to map the space Y onto the circle by $f(s) = (\cos 2\pi s, \sin 2\pi s)$ and $f(\spadesuit) = (1, 0)$. This is clearly one-one and onto. Arc length is a distance on the circle. Arc length means distance between two points is defined as the angle they make with the origin; you take the one that is at most π. This distance gives you the usual notion of convergence on the circle.

Now argue that the map f is nearly an isometry; well it is not, it multiplies distance exactly by 2π.

The essential point of this exercise is the following. In the space Y, if you take a point x different from $\spadesuit$, then a sequence converges to x iff it converges in the space X already. Further, in the space Y, a sequence of points, different from $\spadesuit$ converges to $\spadesuit$ iff they converges to either zero or one in the space X.

2. det one matrices:

The problem is to show that the space of matrices of determinant one is connected. We fix a k and consider only matrices of order $k \times k$. Recall the concept of convergence is entry wise; that is, a sequence of matrices M_n converges to a matrix iff for each i, j their (i, j)-th entries converge. This is same as identifying the space with an appropriate subset of Euclidean space of dimension $k(k-1)/2$.

The plan is to show that it is path connected. Probably you can do easily using the Lie group formalism, or simple group theory and connected components, I have not thought about it. Probably, you can also do it by any one of several canonical reductions, like echelon form etc. But there are several hands on calculations, via: orthogonalize, normalize, (Gram-Schmidt) and rotate — go to identity matrix from any where. This is lengthy but you can easily walk through. Another one is to go to signed permutation matrix and then go to identity. I shall outline the later.

In what follows, path always means continuous path. Continuity is easy for you to verify and so we do not mention. Remember Image of a continuous function defined on an interval is a connected set. Thus if you exhibit a path from a given matrix (of the space) to identity matrix, then the space is connected. Because then, there is a path between any two matrices. In case there is a disconnection of the space as $A \cup B$, union of two disjoint non-empty closed sets; then a path from any point of A to any point of B gets disconnected.

So let M be a given $k \times k$ matrix of determinant one. We shall show a path from it to the identity matrix. this is done in two steps. Let $e_1, e_2, \ldots, e_k$ be the standard orthonormal basis. We show a path from M to the matrix where each row is $\pm e_i$ and all i appear. Second step is to show a path from there to I.

Just keep in mind that non-singularity means just that the rows are independent. Determinant is a continuous function of the matrix. Thus rows of M are independent and span all of k-dimensional space R^k.

Let r_1 be the first row of M. There is one e_i which is not in the span of the other rows. Fix one such. Let $r_i = ae_i + v$ where v is in the span of the other vectors and more importantly, $a \neq 0$. Let $f(\lambda)$ be the matrix with first row $ae_i + \lambda v$ and other rows as they are. No matter what λ is, this is not in the span of the vectors $\{r_2, r_3, \cdots, r_n\}$ Because if it were then ae_i would be and $a \neq 0$ tells e_i would be. Thus independence of rows of M and the fact that this new first row is not in the span of the rest tells you that rows of $f(\lambda)$ are also linearly independent and hence this is non-singular. If g is f divided by its determinant (divide one row); so that det remains one; you have a path from the given matrix to a matrix whose first row is ae_i with $a \neq 0$. All this is to make sure that your path lies in your space.

Now consider the matrix with first row λe_i where λ ranges between a and 1 in case $a > 0$; and ranges between a and -1 in case $a < 0$(and divide by det). You get a path to the matrix whose first row is $\pm e_i$. Remember i and the sign $\pm$ is not at our disposal, depends on the given matrix.

Thus given matrix has a path to a matrix with first row $\pm e_i$ for an i and the other rows remain as they are. Most important, the path is in our space. Now look at the second row and span of the other $(k-1)$ rows and argue exactly as above. Remember $\pm e_i$ is already in the first row now, the new vector you capture is some $\pm e_j$ with $j \neq i$. Get a path to a matrix whose first two rows are $\pm e_i, \pm e_j$.

Continue by induction to get a path from the given matrix to a matrix whose rows are $\pm e_1, \pm e_2, \cdots, \pm e_k$ in some order. Of course if these rows are $e_1, \cdots, e_k$ then we have I. But we do not know. Instead of going from here, we shall come from I to this matrix with a path. We start with two simple observations.

Consider only 2×2 matrices. There is a path from the matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

to I. In fact the path is

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

from $\theta = 0$ to $\theta = \pi$ (may be in the reverse direction). Note that all the matrices above are det one matrices (rotation).

Second observation is the following; again only 2×2 matrices. There is a path from

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

to I. In fact the the same path as above, but now from $\theta = 0$ to $\theta = \pi/2$. similarly, there is a path from

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

to I.

Now let us return to our original problem. We start with I and make a path to a matrix A of det one whose rows are $\pm e_i$. Just to make writing simpler, let us denote f_i is $+e_i$ if it appears in the matrix A; f_i is $-e_i$ if this appears in A. Thus f_i appears as arrow in A and this is $\pm e_i$.

Step 1: A consists of rows $f_1, f_2, \cdots, f_k$ in that order.

Thus A is diagonal and -1 appears at an even number of places, because det is one. Thus take two places and in that coordinate axes apply the rotation mentioned above to I to make those negative. Repeat until you achieve the matrix A. Since -1 is at even number of places, changing two at a time yields final result.

Step 2: Suppose we used e_i in the j-th row and $-e_j$ at i-th row and others are e_p at p-th row $(p \neq i, p \neq j)$. Or we have used $-e_i$ in the j-th row and e_j at i-th row and others are e_p at p-th row $(p \neq i, p \neq j)$. Thus one negative sign and one row permutation applied to I, so that the determinant remains one. In both cases the rotations in the corresponding coordinate axes will give the desired result.

Step 3: general case. Since A consists of rows f_i in some order, there is a permutation π of $\{1, 2, \cdots, k\}$ such that $f_{\pi(i)}$ appears in the i-th row. Any permutation is a composition of transpositions and each transposition changes the sign of the determinant. Using these two facts and the rotations in two coordinate axes at a time we can complete the proof.

The main point is either you have even number of minus signs and even number of transpositions OR odd number of minus and odd number of transpositions.

This can be made precise and something needs to be written, but I am not writing because I believe there must be a better way. Probably you show that det positive matrices is connected and our space is an obvious continuous image of it.

3. Baire category and polynomials:

Let f be an infinitely differentiable function on $[0, 1]$.

$(\exists\ k)\ (\forall\ x)\ f^{(k)}(x) = 0$ implies f is a polynomial.

In fact it is a polynomial of degree at most k that appears in the existential quantifier. This you already know.

It is a simple 'backward' argument. Take such k. since derivative of $f^{(k-1)}$ is zero conclude $f^{(k-1)}$ is a constant a_0. conclude $f^{(k-2)}$ is $a_0 + a_1 x$. Then conclude $f^{(k-3)}$ must be $a_0 + a_1 x + a_2 x^2$ etc.

Of course instead of etc you use induction. We have done it.

The problem to be settled is

$(\forall\ x)\ (\exists\ k)\ f^{(k)}(x) = 0$ implies f is a polynomial.

Thus for each x there is some number $n(x)$ such that the $n(x)$-th derivative at x equals zero.

I had already mentioned that this is not an easy problem. Actually the idea is simple. Let us start with some observations. The first two observations spell out the idea and the rest is just an elaboration using bare.

1^o. Suppose we know $f^{(5)} = 0$ on $(1/3, 2/3)$ and $f^{(8)} = 0$ on $(1/4, 1/2)$. Then actually $f^{(5)} = 0$ on the union $(1/4, 2/3)$.

This is because f is a polynomial $P(x)$ in $(1/3, 2/3)$ and a polynomial $Q(x)$ in $(1/4, 1/2)$. Thus in the common part f equals both P and Q. In other words the two polynomials equal on the common part, which is a non-degenerate interval. This implies that the two polynomials must be same.

2^o. Suppose (i) at every point x of C Cantor set $f^{(8)}(x) = 0$ and (ii) f in each open interval complementary to C, we know f is a polynomial (depending on the interval). Thus in one deleted interval it may be poly of degree at most 20 and in other deleted interval it may be poly of degree at most 100. Then we claim that actually f is a polynomial of degree at most 8.

You see the interesting formulation. If $f^{(8)}(x)$ is zero for all points of an open interval you can say that f is a poly of degree at most 8 on that open interval. But if $f^{(8)}(x)$ is zero for all points of Cantor set, it does not make sense to say that f is a poly of degree at most 8 at all points of the Cantor set.

To get clear idea whether such a thing is possible at all, take the function $h(x) = d(x, C)$ which is continuous and is zero exactly at points of C. (We can give many examples of such continuous functions which are zero exactly at points of C). If

$$f(x) = \int_0^x h(y) dy$$

then derivative of f is zero exactly for points of the Cantor set. You can keep taking indefinite integrals. pause and think.

Let us prove the assertion 2^o. The proof needs a couple of steps.

First we claim that not only $f^{(8)}(x)$ but $f^{(k)}(x)$ is zero for all $k \geq 8$ and all $x \in C$. The crucial point is that f is differentiable any number of times and every point of C is a limit point of C. Given $x \in C$, the second property enables you to take $p_n \in C$, $p_n \neq x$ and $p_n \to x$. The first property enables you to calculate

$$f^{(9)}(x) = \lim \frac{f^{(8)}(p_n) - f^{(8)}(x)}{p_n - x} = 0.$$

Thus $f^{(9)}(x) = 0$ for all $x \in C$. Now use induction.

The plan is to show that at every point $f^{(8)}(x) = 0$. Of course this is granted to you for points in C. Need to show for points outside C, that is, in each of the deleted intervals.

So let us fix any interval (a, b) complementary to C. Thus a and b are in C. Let us say f is a poly of degree k on this interval. In case $k \leq 8$, then there is nothing for you to do; by differentiation, you get $f^{(8)}(x) = 0$ on this interval. Suppose $k > 8$. For instance say $k = 9$. Then $f^{(8)}(x)$ is a continuous function on $[a, b]$ and equals zero at end points and its derivative, namely $f^{(9)}$ is zero inside the interval should tell you, by integration, $f^{(8)} \equiv 0$ on all of $[a, b]$.

In the general case you repeat the same argument. Suppose $k > 8$. Remember: we know higher derivatives are all zero at points of C. Thus $f^{(k-1)}$ is zero at the end points a and b of the interval $[a, b]$ and it is continuous and $f^{(k)}$ is zero inside. Thus $f^{(k-1)} \equiv 0$ on $[a, b]$. Now argue $f^{(k-2)} \equiv 0$ on $[a, b]$ and continue till you reach 8. This completes the proof of 2^o.

3^0. Let us return to our problem. Set

$$A_k = \{x : f^{(k)}(x) = 0\} \qquad k =, 1, 2, 3, \cdots.$$

Then each A_k is closed because each derivative is continuous and $\cup A_k = [0, 1]$ by hypothesis.

The sets A_k, as of now, may neither increase nor decrease.

For example, $f(x) = x^3 + x$ has $f'(0) \neq 0$ but $f''(0) = 0$.

For $g(x) = x^2$ we have $g'(0) = 0$ but $g''(0) \neq 0$.

However, their interiors A_k^o are increasing. This is because, if the k-th derivative is zero on an open interval, then the later derivatives are zero on that open interval.

But interestingly, suppose A_k^0 is disjoint union of non-empty open intervals $I_p = (a_p, b_p) : p = 1, 2, \cdots$. Then none of these end points can be interior points at any time later. This is precisely the content of observation 1^o. Thus any future interiors that get added are disjoint to what you already have.

This observation has the following implication. Let

$$V = \cup A_k^o$$

say V is the disjoint union of open intervals $I_n : n \geq 1$. Remember any (non-empty) open set is union of disjoint non-empty open intervals in only one way. Then each of these I_n is already contained in some A_k^o.

This in turn has the following implication. In each I_n our f is a polynomial.

Let us denote $P = [0,1] - V$. There are two possibilities: either P is empty or not. Suppose it is empty, then V is all of $[0,1]$. In other words the increasing union A_k^o covers all of the compact $[0,1]$ and so must already equal this for some k and this then completes what we wanted to set out to prove.

Let us now assume that $P \neq \emptyset$. We claim that it must be a perfect set. Indeed it is closed being complement of open set V. If it has a point which is not a limit point, then, we have two intervals (a,b) and (b,c) in V but b is not. But then as mentioned above $f^{(i)} \equiv 0$ in (a,b) and $f^{(j)} \equiv 0$ in (b,c). If $m = i \vee j$ then $f^{(m)} = 0$ on both the intervals and hence, by continuity, on all of (a,c). Contradiction because then $(a,c) \subset A_m^o$.

Thus P is a non-empty perfect set. Of course

$$P = (A_1 \cap P) \ \cup \ (A_2 \cap P) \ \cap \ (A_3 \cap P) \ \cup \cdots.$$

But P is a polish space in its own right and these are closed subsets of P and thus one of them must contain an open subset of P. In other words there is an k and open interval (a,b) such that

$$(a,b) \cap P \subset A_k \cap P$$

. Now argument 2^o leads to a contradiction. Every point of P in (a,b) is a limit point of P and so not only $f^{(k)}(x) = 0$ for all $x \in (a,b) \cap P$ we actually have

$$f^{(m)}(x) = 0; \quad \forall m \geq k \quad \forall x \in (a,b) \cap P.$$

This will help you to show (either by differentiation or by integration, as the case may be) to show $f^{(k)}(x) = 0$ for all x in any of the deleted intervals. This ultimately shows that $f^{(k)}(x) = 0$ for all $x \in (a,b)$. In other words $(a,b) \subset A_k^o$ and would have no point in P. This is a contradiction.