Last week we discussed that the three problems — the geometric problem of drawing tangent to a curve at a given point on the curve, the mechanical problem of understanding the rate at which a particle in motion is travelling at a particular instant, the problem of closely approximating a function near a point by a straight line — all lead to the concept of derivative. We saw some basic properties of derivatives.

To get a feel, let us look at an instructive example. Consider the function $f(x) = \sin(1/x)$ defined for non-zero real numbers. It oscillates so badly that we can not assign a value for the function at $x = 0$ so that it is a conitnuous function on R.

Consider the function $f(x) = x\sin(1/x)$ again defined for non-zero real numbers. If we define $f(0) = 0$ then the function is a continuous function on R. However, it is not differentiable at zero. Of course, it is differentiable at all other points.

Consider the function $f(x) = x^2\sin(1/x)$ defined for $x \neq 0$. If we declare $f(0) = 0$, it is a continuous function on R. It is now differentiable also at the point $x = 0$ and in fact $f'(0) = 0$. As already noted, it is differentiable at all non-zero points. Thus f is differentiable at all points and we have the function f' on all of R given by

$$f'(x) = \begin{cases} -\cos(1/x) + 2x\sin(1/x) & \text{if} \quad x \neq 0 \\ 0 & \text{if} \quad x = 0 \end{cases}$$

Whne x is near zero the first term oscillates and second term is near zero. Thus you can see that f' is not continuous at the point zero.

Consider the function $f(x) = x^3\sin(1/x)$ defined for $x \neq 0$ and we declare $f(0) = 0$. Then you can see, f is differentiable and f' is defined on all of R and is a continuous function.

These are pointed out so that you can see how we are achieving more and more 'smoothness'. We shall now see how derivative of f helps us to understand the function f itself better.

Definition; Let f be a function defined on a set S and $a \in S$. We say f has a local maximum at the point a, if there is an open interval (α, β) containing the point a such that $f(x) \leq f(a)$ whenever $x \in S$ and $\alpha < x < \beta$. In other words, locally, that is, in the interval (α, β) the function f has maximum at the point a. Of course, there may be points x outside this interval which are in S and $f(x) > f(a)$. If $a \in S$ is a such a pont that $f(x) \leq f(a)$ for every x in S, then we say that a is a point of global maximum. It is local maximum, but more than that.

We have similar definition regarding minimum. We say f has a local minimum at the point a, if there is an open interval (α, β) containing the point a such that $f(a) \leq f(x)$ whenever $x \in S$ and $\alpha < x < \beta$. Thus locally, that is, in the interval (α, β) the function f has minimum at the point a. There may exist points $x \in S$ such that $x \leq \alpha$ or $x \geq \beta$ where $f(x) < f(a)$. If $a \in S$ is a such a pont that $f(a) \leq f(x)$ for every x in S, then we say that a is a point of global minimum.

Fact: Let f be dfined on an interval (u, v) and $a \in (u, v)$ is a local maximum or loal minimum. If f is differentiable at the point a then $f'(a) = 0$.

Proof is simple. suppose a is a local maximum. Let us take $a \in (\alpha, \beta) \subset (u, v)$ so that $f(x) \leq f(a)$ for all $x \in (\alpha, \beta)$. if you take a sequence of points $\{x_n\}$ in this interval so that each $x_n < a$ and $x_n \uparrow a$, then we see

$$f'(a) = \lim \frac{f(x_n) - f(a)}{x_n - a} \geq 0.$$

similarly, if we take a sequence of points $y_n \downarrow a$ in this interval, each $y_n > a$, we see

$$f'(a) = \lim \frac{f(y_n) - f(a)}{y_n - a} \leq 0.$$

hence $f'(a) = 0$. similar proof applies for local minimum.

The main point is the following. if you are looking for the (local) maximum or local minimum in an open interval, just search among points where the derivative vanishes (assuming that the function is differentiable). Of course you will ask,how do I know if we have max or min? The answer has to wait till we define second derivative.

Understand carefully what we said. a is a a point of max or min implies $f'(a) = 0$. We did not say: $f'(a) = 0$ implies a is a point of max or min. if you consider $f(x) = x^3$, we see $f'(0) = 0$ but zero is neither max nor min.

We also said that the point of max or min must be inside the interval, not endpoint. For example $f(x) = x$ defined on $[3, 20]$ has minimum and maximum at end points and the derivative there exists and not zero. Of couse, you should consider left derivative at the point 3 and left derivative at the point 20.

Suppose that f is a continuous function on a closed bounded interval $[u, v]$. we do know that f is bounded and attains the bounds. suppose a is a point where f assumes largest value. If we now assume that $u < a < v$ and f is differentiable at a we can conclude $f'(a) = 0$. Gometrically it says the following. The the tangent at a is parallel to the x-axis, its slope is zero.

We shall now show that there is a tangent parallel to any given chord of the graph of f.
Fact: Let $u < v \in R$. Let f be a continuous function on an interval $[u, v]$ which is differentiable at every point $u < x < v$. Then there is a point $\theta \in (u, v)$ such that

$$f(v) - f(u) = (v - u)f'(\theta).$$

You can rewrite this as

$$f'(\theta) = \frac{f(v) - f(u)}{v - u}.$$

Observe that slope of the chord joining the two points $(u, f(u))$ and $(v, f(v))$ on the curve (graph of f) is precisely the right side above. The slope of the tangent at the point $\theta \in (u, v)$ is precisely the left side above.

Proof is simple. Shall convert the problem to one we already solved. define

$$\varphi(x) = [f(v) - f(u)]x - [v - u]f(x).$$

then $\varphi(u) = uf(v) - vf(u) = \varphi(v)$ and φ is differentiable in (u, v). Also $\varphi'(x) = [f(v) - f(u)] - [v - u]f'(x)$. If φ is constant then any point $\theta \in (u, v)$ will do the job. Otherwise φ must take values either larger than the value at the end points or values smaller than at the end points. Thus either maix-mum or minimum must be attained in the open interval (u, v). Such a point will do, by the previous fact.

This theorem is called *mean value theorem.* $[f(v) - f(u)]/[v - u]$ is the ‘mean velocity’ in this interval — ratio of distance travelled to the time duration. There is another interpretation. For ten numbers their mean value or average value is their sum/10. If you have a bounded function g on a finite interval, then its mean value or average value is integral/length, integral over the interval and length of the inteval. If you take $g = f'$, defined on (u, v), then its mean valus is precisely $[f(v) - f(u)]/[v - u]$. Thus the theorem says that the mean value of f' is actually value of f' at one of the points in the interval. Note that this quantity is not the mean of the values of the function f at the two end points.

Fact: Let f be defined in an open interval (u, v) differentiable at every point.

(i) If $f'(x) \geq 0$ for all x then f is increasing function in this interval.
(ii) If $f'(x) \leq 0$ for all x then f is decreasing function in this interval.
(iii) If $f'(x) = 0$ for all x then f is a constant in this interval.

Proof is simple. (i) Let $x < y$, then for some point θ,

$$f(y) - f(x) = (y - x)f'(\theta) \geq 0.$$

Same applies to prove (ii). (iii) follows from (i) and (ii).

polynomials of infinite degree:

We have been using exponential function, sine and cosine functions. We have defined exponential function as a sum of series of powers of x and identified it with the power e^x by using continuity and the functional equation $e(x)e(y) = e(x + y)$. This last equation is a consequence of Cauchy product theorem of series. We have used sine and cosine functions and in fact used their derivatives too. So the question is: are we depending on our high school definition here? If so what is it? Did we prove that it is continuous and differentiable?

Let us, once and for all, prove a general theorem. This theorem allows us to cook up functions, show their continuity, differentiability etc. This thorem also allows us to calculate their derivative by providing a formula, just like the one you have for polynomials.

The functions we are talking about are polynomials of infinite degree, sounds like an oxymoron, because when we say polynomial we mean a finite sum of powers of x and thus polynomial can not be of infinite degree. Yes, true these functions are not really clled so but are called power series.

An expression of the following form where α_i, are real numbers is called a power series.

$$P(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \cdots\cdots$$

This is only a symbolic expression. At this stage there is no meaning to it. We shall now give a meaning to it. It is nothing new. Take a real number and substitute that number for x in the above series. You get a series of numbers. This series may converge or may not. if it converges then we associate that value, namely sum of the series as the meaning of the above series at the point x ypu have taken.

Thus a value is attached for the above series whenever you take a real number and find that the series converges. To put it differently, we have a function defined on the set S of all numbers a such that the series converges when you substitute the number a for x in the above series. The important question is: what exactly is the nature of the set S and how does this function behave on that set.

For example can the set S be, say, the interval $[4, 5]$ or union of two intervals like $[1, 2] \cup [9, 10]$? the amazing thing is that such a situation can not occur. If the seris converges when you put $x = 5$, then it converges for every value $|x| < 5$. In fact whenever you take a number a with $|a| < 5$, then the series $\sum |\alpha_n a^n|$ converges, that is, the series $P(a)$ converges absolutely. Be careful, we did not say that the series defining $P(5)$ itself converges absolutely. We did not even say that the series obtained by putting $x = -5$ converges. If the series does not converge when $x = 5$, then it does not converge for any value larger than 5. Here is the main theorem about the power series.

Theorem: Let

$$P(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \cdots\cdots$$

be a power series. Let

$$r = \frac{1}{\limsup \sqrt[n]{|\alpha_n|}}$$

We take $r = 0$ in case the above limsup is ∞. We take $r = \infty$ if the above limsup is zero.

(i) The series converges absolutely for any value x with $|x| < r$. The series does not converge for any value x with $|x| > r$. When $x = \pm r$, it may or may not converge. r is called the radius of convergence of the power series.

Let us assume $r > 0$. and define the function

$$P(x) = \sum_{i \geq 0} \alpha_i x^i; \quad x \in (-r, r).$$

(ii) P is a continuous function on the interval $(-r, r)$.
Let $Q(x)$ be the power series

$$Q(x) = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 + 4\alpha_4 x^3 + \cdots.$$

(iii) The power series Q also has the same radius of convergence r. Thus

$$Q(x) = \sum_{i \geq 0} (i+1)\alpha_i x^i; \qquad x \in (-r, r)$$

is a well defined function.

(iv) The function P is differentiable in the interval $(-r, r)$ and $P'(x) = Q(x)$.

This theorem has a long proof. We shall develop some tools needed as we go along. But before that, a few comments.

First you should realize that the content of the theorem is trivial: you can treat this infinite degree polynomial as if it is usual polynomial you came across in high school. You can differentiate term by term, as you were doing with polynomials in high school. The only difference is that a polynomial is defined for every $x \in R$. But the power series is defined only on an symmetric interval around zero (which, of course, may turn out to be all of R in some cases depending on the numbers α).

Secondly, we have denoted the symbolic infinite series by $P(x)$. We have used the same symbol to denote the function defined on the interval $(-r, r)$. This is done just to avoid too many notational symbols. If you do not like, you can denote the function by f. Thus $P(x)$ stands for the power series $\sum \alpha_i x^i$ without any explanation as to what the symbol x is, what the meaning of the sum is, whether it exists etc. On the other hand, $f(x)$ stands for the function defined on the interval $(-r, r)$ whose value at a point a in this interval is given by the sum $\sum \alpha_i a^i$. But it is not necessary to make such a fine distinction unless you are getting confused.

Thirdly, you must be wondering about the word 'radius' in naming r as the radius of convergence. You can ignore this terminlogy and just think of the interval $(-r, r)$ as interval of convergence. But the story is different. You can think of even putting a complex number like $1+i$ for x in the power series $P(x)$ and want to know whether the series of complex numbers so obtained converges. The answer is the following. Draw a circle of radius r in the (x, y)-plane with origin as center. if your complex number z is inside this circle, then the series $P(z)$ converges. If your complex number z is outside the circle, then the series $P(z)$ does not converge. For points z on the circle, the series $P(z)$ may or may not converge. However, we shall not peep into complex numbers now.

Finally, we assumed that $r > 0$. Because if the series converges only for $r = 0$, then the function is not defined on an interval and the question of coninuity is simple and the question of derivative does not make sense (why?).

underlineProof of (i): Actually we have already done this. Let $\sum_n a_n$ be a series. we have proved the following:

$$\limsup \sqrt[n]{|a_n|} < 1 \Rightarrow \sum |a_n| \text{ converges;}$$

$$\limsup \sqrt[n]{|a_n|} > 1 \Rightarrow \sum a_n \text{ does not converge.}$$

Fix a number x. The series $\sum \alpha_n x^n$ converges if $\limsup \sqrt[n]{|\alpha_n} \, |x| < 1$. Thus if the limsup here is zero, then the series converges for every x. If the limsup is ∞, then this does not converge for any non-zero x. if the limsup is finite and non-zero, then the series converges if $|x| < r$. In this case the series converges absolutely.

On the other hand if $\limsup \sqrt[n]{|\alpha_n|}|x| > 1$, that is, if $|x| > r$, then the seies does not converge. This completes proof of (i).

<u>Towards Proof of (ii):</u>

Let $P_n(x) = \sum\limits_0^n \alpha_i x^i$. Then,by definition of sum of series, we see that for each $x \in (-r, r)$, the sequence $\{P_n(x)\}$ converges to $P(x)$. We also see that for each fixed n the function $P_n(x)$ is a polynomial and is hence continuous. Unfortunately, in general, such a point-wise a limit of a sequence of continuous functions need not be continuous.

For instance, consider the function $f_n(x) = x^n$ defined on the interval $[0, 1]$ for each $n = 1, 2, \cdots$. Let $f(x) = 0$ for $0 \leq x < 1$ and $f(1) = 1$. Clearly, $f_n(x) \to f(x)$ for each $x \in [0, 1]$. Each of the functions f_n is a continuous function on $[0, 1]$ but yet f is not continuous.

When we say $f_n(x) \to f(x)$ for every x what do we mean? suppose we fix a point x, then $f_n(x) \to f(x)$, that is, given $\epsilon > 0$, there is an n_0 such that $n \geq n_0$ implies $|f_n(x) - f(x)| < \epsilon$. This n_0 depends not only on the given ϵ but also on the point x fixed. In the above example, this is what is happening. If $x = 1/2$ you see $f_n(1/2) < 1/1000$ already for $n \geq 10$. But if you want $f(0.999) < 1/1000$ then this $n_0 = 10$ will not do, you need to take much larger n_0. For each point $x < 1$, we do have $f_n(x) \to 0$, but how long should we wait to be close to zero depends on the point x.

If we can make f_n close to f irrespective of the point, then f will be continuous. Let us make it precise. suppose we have a sequence of functions $\{f_n\}$ and a function f, all defined on a set S. Let us say that $f_n \to f$ **uniformly** if given $\epsilon > 0$, there is an n_0 such that $|f_n(x) - f(x)| < \epsilon$ whenever $n \geq n_0$ whatever be the point $x \in S$. Observe that this implies, in particular, that for every point x we have $f_n(x) \to f(x)$. Uniform convergence is more than simply saying point-wise convergence.

Temporarily take $S = [0, 1]$ so that all our functions are defined on $[0, 1]$. Imagine the graph of f and the graphs of $f + \epsilon$ and $f - \epsilon$. here $f \pm \epsilon$ means the function $f(x) \pm \epsilon$. The graph of $f + \epsilon$ is parallel to the graph of f and is above graph of f. Draw pictures. The graph of $f - \epsilon$ is again parallel to graph of f and is below graph of f. If $f_n \to f$ uniformly, then after some stage, graphs of all the functions f_n lie within this band $f - \epsilon$ to $f + \epsilon$. Just like an interval $(a - \epsilon, a + \epsilon)$ in the real line, we can imagine a band of functions, namely, all functions whose graphs lie between $f - \epsilon$ and $f + \epsilon$. Just as convergence of numbers, $x_n \to a$, demands that after some stage all numbers x_n lie within $(a - \epsilon, a + \epsilon)$; uniform convergence of functions, $f_n \to f$ uniformly, demands that after some stage all the functions should lie within the band $(f - \epsilon, f + \epsilon)$.

Convince yourself that in the above example of the sequence of functions $\{x^n\}$ on the interval $[0, 1]$ the convergence is not uniform. if we were to consider the same sequence of functions on the set $S = [0, 0.99]$ the convergence is indeed uniform. if you want to make $|f_n| < \epsilon$, choose k so that $(0.99)^k < \epsilon$. then whatever be $n > k$ and whatever be $x \in S$, we see that $x^n < \epsilon$.

We now make a very useful observation:

Fact: If $f_n \to f$ uniformly on S and if each f_n is continuous on S, then f is continuous on S.

Proof is simple. Suppose that $a \in S$ and $\epsilon > 0$ are given. We produce

$\delta > 0$ so that $|f(x) - f(a)| < \epsilon$ whenever $|x - a| < \delta$ and $x \in S$. Idea is this: $f(x)$ is close to $f_k(x)$ (if k is sufficiently large); and $f_k(x)$ is close to $f_k(a)$ (by continuity of f_k) and $f_k(a)$ is close to $f(a)$. Here is the execution.

First you choose k so that $|f_k(y) - f(y)| < \epsilon/3$ for every $y \in S$. You can do this because of uniform convergence. Actually uniform convergence tells you that you can choose an k so that this inequality is true for every $n \geq k$. But we need not bother about it now. That the inequality holds for this one k is enough. Since f_k is continuous at the point $a \in S$, choose $\delta > 0$ so that $|f_k(x) - f_k(a)| < \epsilon/3$ whenever $|x - a| < \delta; x \in S$. Now if you take any $x \in S$ with $|x - a| < \delta$ we have

$$|f(x) - f(a)| \leq |f(x) - f_k(x)| + |f_k(x) - f_k(a)| + |f_k(a) - f(a)| < \epsilon.$$

returning to our power series, if only we knew that $P_n \to P$ uniformly on $(-r, r)$, we could have immediately applied the above theorem rightaway. Since each P_n, a polynomial is continuous, it follows that P is also continuous. Unfortunately, there is a little twist. In general P_n does not converge uniformly on all of $(-r, r)$. They nearly do so.

Fact: Let $0 < c < r$. Then $P_n \to P$ uniformly on $[-c, c]$.

You can take $c = r - 0.00001$, but can not take $c = r$ in general.

Proof is simple. Let $\epsilon > 0$. Remember the power series converges absolutely at the point c. That is, the series $\sum |\alpha_i| c^i$ is convergent. Also remember that when a series converges, the partial sums converge to the grand sum. A consequence of this, which we observed earlier , is that tail sums converge to zero. That is if $t_n = \sum\limits_{i>n} |\alpha_i| c^i$, then $t_n \to 0$. Let us choose k so that $t_k < \epsilon$. We claim that $|P_n(x) - P(x)| < \epsilon$ for any $n > k$ and for any $x \in [-c, c]$. In fact for any such n and x,

$$|P(x) - P_n(x)| = |\sum_{i>n} \alpha_i x^i| \leq \sum_{i>n} |\alpha_i| c^i \leq \sum_{i>k} |\alpha_i| c^i < \epsilon.$$

We now complete proof of (ii).

Let us take $a \in (-r, r)$. We show that P is continuous at a. First fix $c > 0$ so that $-r < -c < a < c < r$. The above two facts tell you that P is continuous on the interval $[-c, c]$. To show that P is continuous at a, take a sequence $x_n \to a$, $-r < x_n < r$. Then after some stage $x_n \in (-c, c)$. By continuity of P on $[-c, c]$ tell us $P(x_n) \to P(a)$.

This completes proof of (ii).

<u>Towards proof of (iii):</u>

Recall

$$Q(x) = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 + 4\alpha_4 x^3 + \cdots$$

We need ot show that Q also has radius of convergence r. Here the coefficient of x^n is $(n+1)\alpha_{n+1}$ so for the radius of convergence of Q we would be involved in $\sqrt[n]{(n+1)|\alpha_{n+1}|}$. On the other hand calculation of r involved $\sqrt[n]{|\alpha_n|}$. We first sort out this mismatch between n and $n+1$. We claim that the series $Q(x)$ converges iff the following series converges.

$$xQ(x) = \alpha_1 x + 2\alpha_2 x^2 + 3\alpha_3 x^3 + 4\alpha_4 x^4 + \cdots$$

In fact let x be any real number. if the series $Q(x)$ converges, then multiplying each term by x we get the second series and hence the second series also converges. conversely, suppose that the second series converges. If $x = 0$, then first series also converges (inspect what is the series when you put $x = 0$). So let $x \neq 0$. But then the first series is obtained from the seond series by dividing each term by x.

As a consequence, the radius of convergence of the power series $Q(x)$ is given by

$$\frac{1}{\limsup \sqrt[n]{n|\alpha_n|}}.$$

Thus to prove (iii) we only need to show that

$$\limsup \left(\sqrt[n]{n}\ \sqrt[n]{|\alpha_n|} \right) = \limsup \sqrt[n]{|\alpha_n|}.$$

Fact: Let $\{a_n\}$ and $\{b_n\}$ be sequences of non-negative numbers; $a_n \to a$ $0 < a < \infty$; $\limsup b_n = b$; $0 \leq b \leq \infty$. Then $\limsup(a_n b_n) = ab$.

if this fact is proved, then we can take $a_n = \sqrt[n]{n}$ and $b_n = \sqrt[n]{|\alpha_n|}$. we had already proved earlier that $a_n \to 1$, so that $a = 1$ and the fact applies to give us what we wanted.

Here is proof of the fact.
Case $b = \infty$. Need to show $\limsup a_n b_n = \infty$. Given any number c, we show that infinitely many of these $a_n b_n$ are larger than c. Since $\limsup b_n = \infty$, infinitely many b_n are larger than $c/(a/2)$; say $b_{n_1}, b_{n_2}, b_{n_3} \cdots$. Since $a_n \to a$, after some stage $a_n > a/2$; say for $n \geq k$. If you look at $a_{n_i} b_{n_i}$ for $n_i > k$

you see that they are all larger than c.

Case $b < \infty$.

Since $\limsup b_n = b$, there is a subsequence converging to b. Say

$$b_{n_1}, b_{n_2}, b_{n_3}, \to b.$$

Since $a_n \to a$, every sunsequence also converges to a. Hence

$$a_{n_1} b_{n_1}, a_{n_2} b_{n_2}, a_{n_3} b_{n_3}, \cdots \to ab$$

Thus ab is a limit point of the sequence $\{a_n b_n\}$ so that $\limsup a_n b_n \geq ab$. If c is any limit point of the sequence $\{a_n b_n\}$, then there is a subsequence

$$a_{n_1} b_{n_1}, a_{n_2} b_{n_2}, a_{n_3} b_{n_3}, \cdots \to c.$$

But

$$a_{n_1}, a_{n_2}, a_{n_3}, \cdots \to a.$$

Since $a \neq 0$, we conclude that

$$b_{n_1}, b_{n_2}, b_{n_3}, \cdots \to \frac{c}{a}.$$

Thus c/a is a limit point of the sequence $\{b_n\}$ and hence must be not larger than its limsup.

$$\frac{c}{a} \leq b; \qquad c \leq ab$$

Thus any limit point of the sequence $\{a_n b_n\}$ issmaller than ab. Since we have already shown that ab is a limit point, we conclude that $\limsup a_n b_n = ab$.

This completes proof of (iii)

Finally, we prove (iv):

The idea is similar to that of (ii). Fix $a \in (-r, r)$. Need to show that $[P(x) - P(a)]/[x - a]$ is close to $Q(a)$when x is close to a. We know that $[P_n(x) - P_n(a)]/[x - a]$ is close to $Q_n(a)$when x is close to a where

$$P_n(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n.$$

$$Q_n(x) = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 + \cdots + n\alpha_n x^{n-1}.$$

If we can show that $[P(x) - P(a)]/[x - a]$ is close to $[P_n(x) - P_n(a)]/[x - a]$ and $Q_n(a)$ is close to $Q(a)$, we are done as earlier in (ii). this is what we execute now.

Let $\epsilon > 0$ be given. We shall exhibit $\delta > 0$ so that $(a-\delta, a+\delta) \subset (-r, r)$ and

$$0 < |x-a| < \delta \Rightarrow \left|\frac{P(x)-P(a)}{x-a} - Q(a)\right| < \epsilon.$$

This will show that P is differentiable and $P'(a) = Q(a)$.

First we fix $c > 0$ so that $-r < -c < a < c < r$. This is possible because $-r < a < r$. Since Q also has radius of convergence r and $0 < c < r$ we conclude that the series $Q(c)$ converges absolutely. That is the series $\sum (i+1)|\alpha_{i+1}|c^i$ is convergent. So as earlier in (ii), its tail sums converge to zero. Fix N such that

$$\sum_{i \geq N} (i+1)|\alpha_{i+1}|c^i < \frac{\epsilon}{3}.$$

We have already defined

$$P_n(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n.$$

$$Q_n(x) = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 + \cdots + n\alpha_n x^{n-1}.$$

Let us also put

$$\widetilde{P_n}(x) = \alpha_{n+1}x^{n+1} + \alpha_{n+2}x^{n+2} + \alpha_{n+3}x^{n+3} + \cdots.$$

$$\widetilde{Q_n}(x) = (n+1)\alpha_{n+1}x^n + (n+2)\alpha_{n+2}x^{n+1} + (n+3)\alpha_{n+3}x^{n+2} + \cdots.$$

so that for each $n \geq 1$ we have

$$P(x) = P_n(x) + \widetilde{P_n}(x); \qquad Q(x) = Q_n(x) + \widetilde{Q_n}(x); \qquad (\diamondsuit)$$

With this notation, choice of N tells us

$$|x| \leq c \Rightarrow |\widetilde{Q_N}(x)| \leq \epsilon/3. \qquad (\spadesuit)$$

Note that by mean value theorem, <u>for $x \in [-c, c]$ we have the following</u>.

$$\left|\frac{\alpha_{N+1}x^{N+1} - \alpha_{N+1}a^{N+1}}{x-a}\right| \leq (N+1)|\alpha_{N+1}|c^N \leq \frac{\epsilon}{3}.$$

$$\left|\frac{\alpha_{N+1}x^{N+1} + \alpha_{N+2}x^{N+2} - \alpha_{N+1}a^{N+1} - \alpha_{N+2}a^{N+2}}{x-a}\right|$$

$$\leq (N+1)|\alpha_{N+1}|c^N + (N+2)|\alpha_{N+2}|c^{N+1}$$
$$\leq \frac{\epsilon}{3}.$$

More generally, for every $k \geq 1$, we have

$$\left| \frac{\displaystyle\sum_{i=N+1}^{N+k} \alpha_i x^i - \sum_{i=N+1}^{N+k} \alpha_i a^i}{x-a} \right| \leq \frac{\epsilon}{3}.$$

Since this is true for every $k \geq 1$ we see by taking limits (over k),

$$-c \leq x \leq c \Rightarrow \left| \frac{\widetilde{P_N}(x) - \widetilde{P_N}(a)}{x-a} \right| \leq \frac{\epsilon}{3}. \qquad (\heartsuit)$$

As noted already and easy to see, derivative of the polynomial P_N is Q_N. Thus we can fix $\delta > 0$ so that $(a-\delta, a+\delta) \subset (-c, c)$ and

$$0 < |x-a| < \delta \Rightarrow \left| \frac{P_N(x) - P_N(a)}{x-a} - Q_N(a) \right| < \frac{\epsilon}{3} \qquad (\clubsuit)$$

<u>For $0 < |x-a| < \delta$ we have the following.</u>

In these string of inequalities below, first use $(\diamondsuit)$ and then $|t+u+v| \leq |t| + |u| + |v|$ and then use $(\clubsuit), (\heartsuit), (\spadesuit)$.

$$\left| \frac{P(x) - P(a)}{x-a} - Q(a) \right|$$

$$= \left| \frac{P_N(x) - P_N(a)}{x-a} + \frac{\widetilde{P_N}(x) - \widetilde{P_N}(a)}{x-a} - Q_N(a) - \widetilde{Q_N}(a) \right|$$

$$\leq \left| \frac{P_N(x) - P_N(a)}{x - a} - Q_N(a) \right| + \left| \frac{\widetilde{P_N}(x) - \widetilde{P_N}(a)}{x - a} \right| + |\widetilde{Q_N}(a)|$$

$$\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

This completes the proof of the theorem.

The proof is rather long but you see that each step is simple.