Rev. Acad. Canar. Cienc., XIII (Núms. 1-2-3), 35-38 (2001) (publicado en Julio de 2002)

A PROOF OF THE INTERMEDIATE VALUE THEOREM ADAPTED TO POLYNOMIALS

J. M. ALMIRA AND M. JIMENEZ

1. Introduction

Rolle (1652-1719) published his famous theorem in 1691 (see [5, p. 504]).

Theorem 1. ***(Rolle)*** *Let us assume that $f \in C[a,b]$ has derivative on (a,b). If $f(a) = f(b) = 0$ then $f'(\theta) = 0$ for some $\theta \in (a,b)$.*

After this, many mathematicians used the following variation of the theorem:

Theorem 2. ***(Mean Value Theorem)*** *Let us assume that $f \in C[a,b]$ has derivative on (a,b). Then there exists a certain value $\theta \in (a,b)$ such that*

$$\frac{f(b)-f(a)}{b-a} = f'(\theta)$$

Lagrange (1736-1813) tried to prove the mean value theorem, but he failed, since his point of view was completely wrong in the sense that he tried to make Calculus in a purely algebraic way (see [1], [3], [4]). The first mathematician which sistematically used the above result was Cauchy (1789-1857), who also 'proved' the following generalization:

Theorem 3. ***(Cauchy's Mean Value Theorem)*** *Let us assume that $f, g \in C[a,b]$ both have derivative on (a,b), and $g(a) \neq g(b)$. Then there exists a certain value $\theta \in (a,b)$ such that*

$$\frac{f(a)-f(b)}{g(a)-g(b)} = \frac{f'(\theta)}{g'(\theta)}$$

A proof of Rolle's theorem is as follows: Firstly we use that f attains an extreme value at some $c \in (a,b)$ (whenever it is not constant). Hence $f'(c) = 0$, since it is a derivable function. On the other hand, if f is constant then f' is identically zero.

Of course, it is easy to note that we have used the existence of extreme points for continuous functions defined on intervals. This is a theorem which was formally proved by Weierstrass (much more after than Rolle's theorem) and in a modern perspective, the key for the proof is that intervals are compact sets. The same can be applied to the famous Bolzano's theorem (that a continuous function which changes the sign must vanish at some point). The conclusion is that the topology of real numbers was not an easy matter and, in fact, was the key for the simplification of many results from the foundations of Analysis.

Cauchy's main objective was to introduce the rigor in the proofs of general properties of continuous functions. He said it was neccesary to give exact proofs for continuous

Research partially supported by Junta de Andalucía, Grupo de Investigación "Aproximación y Métodos Numéricos" FQM-0178 and PB-94; email: jmalmira@ujaen.es, mjimenez@ujaen.es.

functions of many results that people know to be true for algebraic functions (see [5, p. 1252]).

We have some doubts with respect to the question if mathematicians of that period had the proofs of some elementary facts of Mathematical Analysis, for the class of real algebraic polynomials. For example: what about the intermediate value theorem? In this note we give proofs adapted for the space of polynomials of some of these results. We believe our proofs could be understood by Cauchy's contemporaries.

2. The proofs

Lemma 4. ***(Increasing Function Theorem, for Polynomials (IFTP))*** *Let us assume that $p(x) \in \mathbb{R}[x]$ satisfies $p'(x) \geq 0$ for all $x \in (a,b)$ ($p'(x) \leq 0$ for all $x \in (a,b)$, respectively). Then p is an increasing function (a decreasing function, respectively) on $[a,b]$.*

Proof. We use the geometric concept of integral to prove that $\int_a^b x^n dx = \frac{b^{n+1}}{n+1} - \frac{a^{n+1}}{n+1}$ (a fact that was already proved by Fermat in 1636, see [5, p. 466]). Hence the relation $p(x) - p(a) = \int_a^x p'(t)dt$ holds trivially true for the class of polynomials. If $p'(t)$ is greater (or equal) than zero on $[a,b]$ then $x_1, x_2 \in [a,b]$, $x_1 \leq x_2$ implies

$$p(x_2) - p(x_1) = \int_{x_1}^{x_2} p'(t)dt \geq 0$$

Hence p is an increasing function in such a case. □

Lemma 5. ***(Intermediate Value Theorem for monotone functions (IVT))*** *Let us assume that $f \in C[a,b]$ is monotone. Then there exist $m, M \in \mathbb{R}$, $m < M$, such that $f([a,b]) = [m,M]$.*

Proof. We assume without loss of generality that f is increasing and $c \in (f(a), f(b))$. Now, we set $A_1 = \{x \in [a,b] : f(x) < c\}$ and $A_2 = \{x \in [a,b] : f(x) > c\}$. They are both non empty bounded sets, so that we can define the quantities $\alpha = \sup A_1$ and $\beta = \inf A_2$. It follows from the monotonicity of f that $\alpha \leq \beta$. Take $\{x_n\}_{n=0}^{\infty} \subset A_1$, $\{y_n\}_{n=0}^{\infty} \subset A_2$ such that $\lim x_n = \alpha$ and $\lim y_n = \beta$. Then $f(\alpha) = \lim f(x_n) \leq c \leq \lim f(y_n) = f(\beta)$, so that $\alpha = \beta$ clearly implies that $f(\alpha) = c$. On the other hand, if $\alpha \neq \beta$ then it is clear that $(\alpha + \beta)/2 \in [a,b] \setminus (A_1 \cup A_2)$, so that $f((\alpha+\beta)/2) = c$. □

We must observe that the notation of supremun, etc, were not known at that time. In fact, people from Cauchy's time were not able to use the axiom of supremun (that we use now easily). In his famous paper, Bolzano used (without proof) the following claim, which is equivalent to the axiom of supremun: *If a property* **P** *is satisfied by all numbers less or equal to a certain number* **x**, *and the property does not hold true for all (real) numbers, then there exists a number, which is the greatest in the set of (real) numbers* **M** *such that* **P** *holds true for all* **y** < **M**. He was never able to prove this property. From our point of view, the proof we have presented here, could be though -but not stated with complete rigor- at that time.

Theorem 6. ***(Intermediate Value Theorem, for Polynomials (IVTP))*** *Let $p(x) \in \mathbb{R}[x]$ be a polynomial with coefficients in $\mathbb{R}$, and let $a, b \in \mathbb{R}$, $a < b$ be fixed. Then there exist $m, M \in \mathbb{R}$, $m < M$, such that $p([a,b]) = [m,M]$.*

Proof. It follows from the algorithm of division of polynomials (Euclid) that a polynomial $p(x) = a_0 + a_1 x + \cdots a_n x^n$ has at most n zeros. This implies that its derivative (which is another polynomial) has also a finite number of zeros (at most $n-1$), so that (using Lemma 4), we can decompose the interval $[a,b]$ as a finite union of closed intervals $[a,b] = I_1 \cup I_2 \cup \ldots \cup I_k$, where $I_s = [t_{s-1}, t_s]$, $a = t_0 < t_1 < \ldots < t_k = b$ and $p_{|I_s}$ is a monotone function for each $s \leq k$. It follows from Lemma 5 that $p(I_s) = J_s = \lfloor m_s, M_s \rfloor$ is a closed interval for all s. Furthermore, $t_s \in I_{s+1} \cap I_s$, so that $J_{s+1} \cap J_s \neq \emptyset$ for all s. This means that $J_{s+1} \cup J_s$ is another closed interval and, repeating the argument several times, we obtain that $p([a,b]) = J_1 \cup J_2 \cup ... \cup J_k = [m, M]$ is a closed interval. □

Of course, there are elementary proofs of the increasing function theorem for derivable functions (see [6]),

Theorem 7. ***(Increasing Function Theorem (IFT))*** *Let us assume that $f \in C[a,b]$ is differentiable on (a,b) and $f'(x) \geq 0$ for all $x \in (a,b)$ ($f'(x) \leq 0$ for all $x \in (a,b)$, respectively). Then f is an increasing function (a decreasing function, respectively) on $[a,b]$,*

but they use Cantor's principle of nested intervals (it asserts that the intersection of a nested sequence of closed intervals whose diameters approach to zero, is a point), which is the key idea for the bisection method (hence for a proof of Bolzano's theorem and, as a consequence, a proof of the intermetiate value theorem for continuous functions).

Another proof of the **IVTP** which could be thought in Cauchy's times is as follows: We can use the Fundamental Theorem of Algebra (firstly proved by Gauss in 1799, see [1]) to decompose the polynomial $p(x) \in \mathbb{R}[x]$ as a product of irreducible factors:

$$p(x) = \prod_{i=1}^{n}(a_i x + b_i)^{n_i} \prod_{j=1}^{m} q_j(x)^{m_j},$$

where the polynomials $q_j(x)$ are quadratic factors without real zeros, so that the sign of $\prod_{j=1}^{m} q_j(x)^{m_j}$ is constant on the real line. Hence the changes of sign in the interval $[a,b]$ are caused by the linear factors $a_i x + b_i$. This clearly implies that Bolzano's Theorem holds for the class of polynomials. The intermediate value theorem is now a corollary of Bolzano's theorem. This proof also serves for polynomials in the field of algebraic real numbers, or any other really closed field. Note that to use this focus, we must firstly prove the fundamental theorem of algebra, without the use of **IVTP**.

Rolle's theorem and the Mean Value Theorem have now a proof for the class of polynomials as corollaries of the **IVTP**. Other consequences are the following theorems:

Theorem 8. *If $p(x) = a_0 + ... + a_{n-1}x^{n-1} + x^n$ is a polynomial of odd degree, then it has a real root.*

Proof. It suffices to note that

$$p\left(1 + \sum_{i=0}^{n-1} |a_i|\right) > 0 > p\left(-\left(1 + \sum_{i=0}^{n-1} |a_i|\right)\right)$$

and use the **IVTP**. □

Theorem 9. *Let $a < b$ be two real numbers and let $p(x)$ be a polynomial with real coefficients. Then p is uniformly continuous on $[a,b]$.*

Proof. Let $[m, M] = p'([a,b])$ and set $C = \max\{|m|, |M|\}$. Then

$$|p(x) - p(y)| = |p'(\theta)||x - y| \leq C|x - y|$$

holds for all $x, y \in [a,b]$. □

Of course, we must recall that many results from a first year Calculus course have trivial proofs for the set of polynomials. For example, the chain rule can be checked by direct computation, since the formula $(q(z)^n)' = (q(z)^{n-1}q(z))' = (n-1)q(z)^{n-2}q'(z)q(z) + q(z)^{n-1}q'(z) = nq(z)^{n-1}q'(z)$, $n = 1, 2, \ldots$ is clear, and

$$(p \circ q)'(z) = \left(\sum_{k=0}^{n} a_k q(z)^k\right)' = \sum_{k=0}^{n} k a_k q(z)^{k-1} q'(z) = p'(q(z))q'(z).$$

We are convinced that some of these elementary proofs where already known at the time of Cauchy. Perhaps he was thinking how to extend these results to the class of continuous functions, when he decided to write his *Cours d'Analyse* [2]. If he was completely right or not, is not the main question. From our point of view, the most important thing is that Cauchy's work had a deep influence in other mathematicians for a long period of time, and it was a great contribution for the introduction of rigor in Mathematical Analysis. The main goal of this note was nothing but to fill the gap caused by Cauchy's claim about the existence of special proofs of some elementary facts from Calculus, adapted to the space of real algebraic polynomials.

References

[1] **Boyer, C.** *A History of Mathematics.* John Wiley & Sons, (1968)

[2] **Cauchy, A. L.** *Cours d'analyse algébrique* (1821), in Ouvres Complètes, Gauthier-Villars, Paris.

[3] **Duran, A. J.** *Historia, con personajes, de los conceptos del cálculo,* Alianza Editorial. Madrid. (1996)

[4] **Grattan-Guinnes, I.** *From the Calculus to Set Theory 1630-1910* Durckworth (1980) (translated into Spanish by Alianza Ed. in (1996)).

[5] **Kline, M.,** *Mathematical Thought from Ancient to Modern Times,* Oxford University Press (1972).

[6] **Tucker, T. W.,** Rethinking rigor in Calculus: the role of the Mean Value Theorem, Amer. Math. Monthly **104** (1997) 231-240.

J. M. Almira, M. Jiménez.
Departamento de Matemáticas. Universidad de Jaén.
E. U. P. Linares, C/ Alfonso X el Sabio, 28.
23700 Linares (Jaén) SPAIN.