University of Cape Town

Department of Mathematics

THE THEORY AND SOME APPLICATIONS OF PTÁK'S METHOD OF NON-DISCRETE MATHEMATICAL INDUCTION

by

T.A. Marsh

A thesis presented in fulfilment of
the requirements for the degree of
Master of Science in Mathematics
at the
University of Cape Town
1977

ACKNOWLEDGEMENTS

I wish to acknowledge with grateful thanks the contributions made by certain individuals to the completion of this thesis. I am indebted to: my supervisor, Associate Professor J.H. Webb, for the many hours he spent teaching, inspiring, guiding and assisting me prior to and throughout the preparation of this manuscript; Associate Professor W. Kotzé for the interest he showed and help he provided during the initial stages of my M.Sc. work; Mrs. Heather Light for her patience, cheerfulness and efficiency in carrying out the typing; my good friend and colleague Mr. Henry Robbins for helpful criticisms and suggestions during the proof-reading stage; and Mr. Gert Gabriels for the excellent quality of his printing.

I acknowledge with gratitude the generosity of Mr. N.M. Paterson (Headmaster of the South African College High School) and the Department of Education of the Cape Provincial Administration for having granted me 18 months of study-leave from my teaching job. I am also grateful for the financial support of a C.P.A. Teaching Bursary. In addition I thank Professor K.O. Househam (Head of the Department of Mathematics of the University of Cape Town) for the interesting and enjoyable experience of working as a Teaching Assistant on his staff.

Finally, I thank my wife Julie for her unfailing tolerance, support and encouragement.

T.A.M.

CONTENTS

SECTION 4 APPLICATION OF THE INDUCTION THEOREM TO PROVIDE EXTENDED CONTRACTION PRINCIPLES

SECTION 5 THE INDUCTION THEOREM APPLIED TO NEWTON'S METHOD

SECTION 6 ASSORTED APPLICATIONS OF THE INDUCTION THEOREM

INTRODUCTION

In functional and numerical analysis, iterative procedures are often devised in order to construct or approximate an element which lies in a given set or satisfies a given relation. At each stage of the iterative process we are dealing with elements which satisfy the desired condition only approximately, the degree of approximation becoming better at each step. For example, to construct a point x belonging to a given set Z we consider a family $Z(\cdot)$ of so-called approximate sets depending on a small positive parameter t. The inclusion $z \in Z(t)$ means that $z \in Z$ is satisfied only approximately, the approximation being measured by the number t. It turns out that, under suitable hypotheses about the behaviour of the function $t \to Z(t)$, it is possible to obtain a convergent iterative process.

V. Pták on analysing the above approximation procedure developed the notion of rate of convergence as a function ω on the domain of t. Using this approach the key issue becomes the possibility of passing from a given approximation $Z(t)$ to a better one $Z(\omega(t))$. This step corresponds to the step from n to $n + 1$ in classical induction proofs. Pták's Induction Theorem, the central idea in his work on non-discrete induction, can in fact be regarded as a continuous analogue of the method of mathematical induction. It gives an abstract model for iterative existence proofs in functional and numerical analysis. It provides a method for the investigation of iterative constructions as well as yielding considerable simplification of proofs. The method involves the construction of a system of relations which describe the desired relation with an increasing degree of accuracy and such that this relation is a limit of the system in a natural manner.

Pták carried out his preliminary investigations in this field during the 1950's and 60's making a detailed study of complete linear topological spaces. His aim was to find ways in which to extend the realms of the Open Mapping and Closed Graph Theorems, using *inter alia* his notions of B-completeness and nearly openness. Eventually he began to delve more deeply into the metric aspects of these two theorems and discovered the essence of their mechanisms to lie in the idea of rates of convergence regarded as functions which he called small functions. Finally, in 1973, he applied the small function approach to the approximation technique underlying the Closed Graph Theorem and reworked this theorem into a more abstract setting. This gave rise to his Induction Theorem.

The ideas involved in this work have been refined and expanded over the past few years and Pták has found several applications of his rather powerful theorem. He has shown that the classical Closed Graph and Banach Fixed Point Theorems are simple consequences of it. In addition he has described applications of the Induction Theorem to the following: an analysis of the rate of convergence of Newton's Method; Moser's Theorem concerning a modification of Newton's Method; the problem of finding eigenvalues of almost decomposable matrices; proof of a Selection Theorem;

factorization of elements of a module over a Banach algebra; Kadison's Theorem concerning strict irreducibility of *-representations of B*-algebras.

The aim of this thesis is three-fold:

(1) to develop the theory of small functions;

(2) to synthesize Pták's work presented in his papers [10], [11],...,[16] into a coherent body of knowledge;

(3) to elaborate on Ptak's work

  (i) by providing small function generalizations of Banach's Fixed Point Theorem and Edelstein's Extended Contraction Principle;

  (ii) by connecting the Induction Theorem to Baire's Category Theorem and Cantor's Intersection Theorem.

Throughout the exposition the editorial "we" is to be understood in the sense of Halmos [18]; "we" means "the author and the reader".

Section 1

RATES OF CONVERGENCE

1.1 Rate of convergence as a function

Instead of regarding the rate of convergence of an iterative process as a number Pták [11] defines it as a function.

1.1.1 **Definition** Let T be an interval of the form $T = \{t : 0 \leqslant t \leqslant t_0\}$ for some number t_0. A *rate of convergence* on T is a function ω with the following properties:

(i) ω maps T into itself;

(ii) for each $t \in T$ the series $t + \omega(t) + \omega(\omega(t)) + \dots$ is convergent.

Such functions ω will be referred to as *small functions*. In the sequel we shall use the abbreviation ω^n for the n-th iterate of ω, so that $\omega^2(t) = \omega(\omega(t))$, etc.

Several examples of small functions can be abstracted from Pták's papers [13], [15] and [16].

1.2 Examples of small functions

1.2.1 Linear small functions

If ω is linear, i.e. $\omega(t) = \alpha t$ for some constant α, then ω is small if and only if $0 \leqslant \alpha < 1$.

1.2.2 Functions of the form $t \to t^a$

If $a > 1$ and $\omega(t) = t^a$, then ω is a small function on the interval $0 \leqslant t < 1$. This is not difficult to show.

Note that $2^{-1/(a-1)} < 1$ and that for $0 \leqslant t \leqslant 2^{-1/(a-1)}$ we have $t^a \leqslant \frac{1}{2}t$. By induction this leads to $t^{a^n} \leqslant \left(\frac{1}{2}\right)^n t$, whence

$$\begin{aligned}\sum_{r=0}^{\infty} \omega^r(t) &= t + t^a + t^{a^2} + \dots \\ &\leqslant t + \tfrac{1}{2}t + \left(\tfrac{1}{2}\right)^2 t + \dots = 2t .\end{aligned}$$

Pták [16] finds this small function particularly useful in his treatment of Moser's Theorem.

1.2.3 A small function derived from the function $z \to \frac{\beta}{\gamma - z}$

If $\gamma^2 > 4\beta$, $\beta > 0$, $\gamma > 0$, the function $z \to \frac{\beta}{\gamma - z}$ gives rise to a small function defined for all $t > 0$:

$$\omega(t) = t\,\frac{\gamma + t - v}{\gamma - t + v} \quad \text{where} \quad v = \sqrt{(\gamma + t)^2 - 4\beta} .$$

This will receive a detailed treatment in Subsection 1.6, while a demonstration of the usefulness of this small function will be given in Subsection 6.1.4 in our discussion of eigenvalues of almost decomposable matrices.

We proceed now to develop the theory of small functions by establishing some of their properties.

1.3 Ways of generating small functions

1.3.1 Comparison with known small functions

Obviously, if ω_1 is small on T and ω_2 defined on T is such that $\omega_2^k(t) \leqslant \omega_1^k(t)$ for all $k \in \mathbb{N}$ and all

$t \in T$, then ω_2 is small on T . The condition $\omega_2(t) \leqslant \omega_1(t)$ for all t is not sufficient as will be shown in Subsection 1.3.5.

1.3.2 A ratio test

$$\text{If } \limsup_{t \to 0^+} \frac{\omega(t)}{t} < 1 \text{ , then}$$

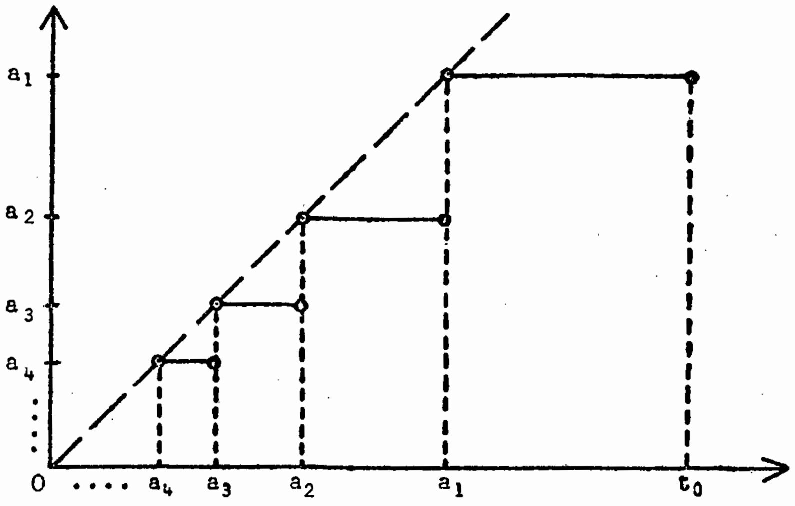
$$\omega \text{ is a small function.}$$

Note, however, that $\omega(t) < t$ for all t is not a sufficient condition for a function to be small as will be shown in Subsection 1.3.4.

1.3.3 Step functions

If $\{a_n\}$ is a sequence such that $\sum a_n < \infty$, $a_n \searrow 0$, then by defining $\omega(t)$ as the first a_n such that $a_n < t$, we generate a small function, as represented in Diagram (1) below. This is a technique we shall make use of in Section 3 in the proof of the Cantor Intersection Theorem.

Diagram 1



1.3.4 A function which, although below $\omega(t) = t$, is not small

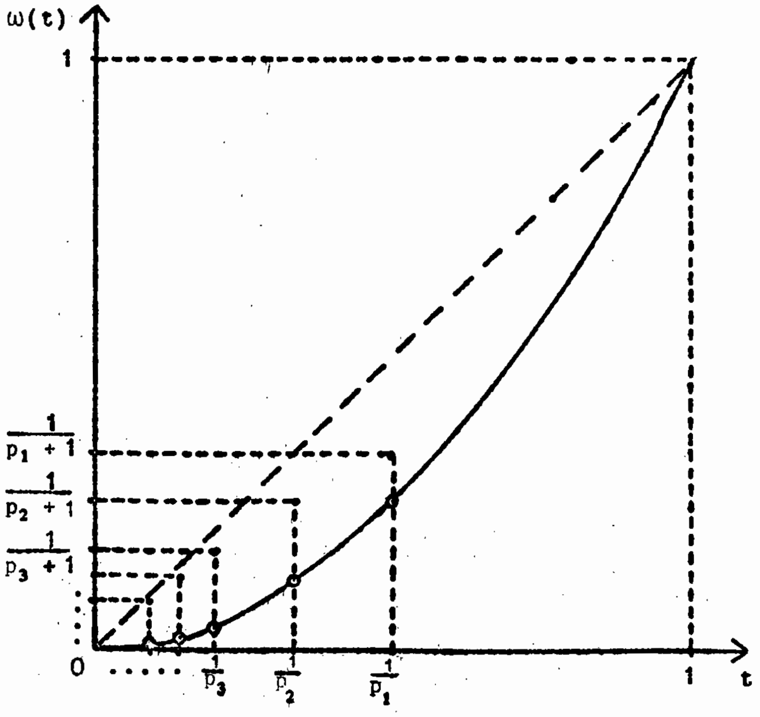
If ω_1 is small and not non-decreasing and ω_1 is levelled out into a step function ω_2, where $\omega_2(t) = \sup\limits_{t_0 \leqslant t} \omega_1(t_0)$ as shown in Diagram (3) below, then ω_2 is not necessarily small.

Let $\{p_n\}$ be the sequence of all prime numbers, T the set $\{t : 0 \leqslant t < 1\}$, and define

$$\begin{cases} \omega_1\left(\dfrac{1}{p_n}\right) = \dfrac{1}{p_n + 1} \\ \omega_1(t) = t^2 \quad \text{if} \quad t \neq \dfrac{1}{p_n} \quad \text{for any} \quad n\ . \end{cases}$$

Note that ω_1 has been chosen such that, if at some stage $\omega_1^i(t) = \dfrac{1}{p_j}$, then at the next iteration $\omega_1^{i+1}\left(\dfrac{1}{p_j}\right) = \dfrac{1}{(p_j + 1)^2}$, and at each stage after this the element of T with which we are concerned will not be of the form $\dfrac{1}{p_n}$ for any n. This function is represented in Diagram (2) below.

Diagram 2

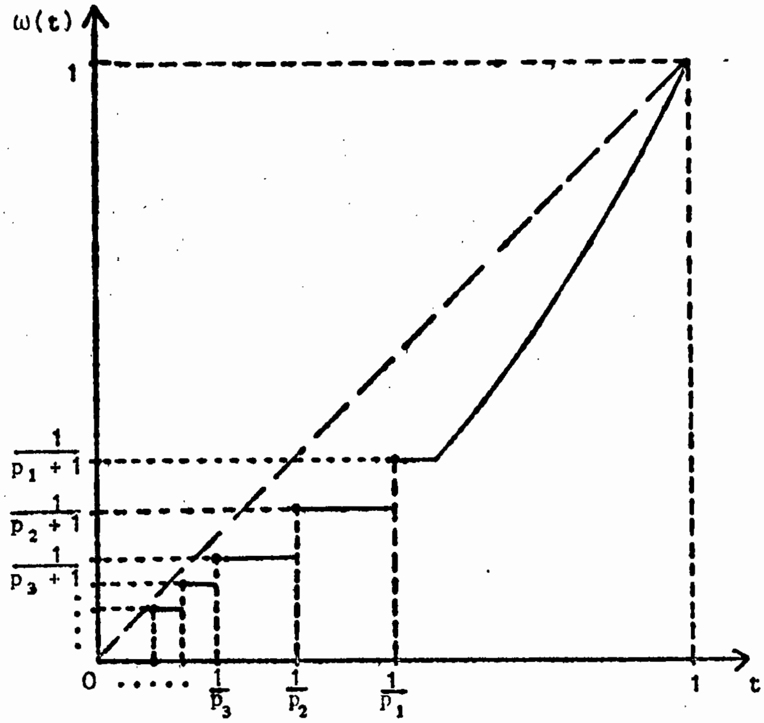


Since $t \to t^2$ is small, as shown in Subsection 1.2.2, it is obvious that $\sum \omega_1^r(t) < \infty$, i.e. that ω_1 is small. But, when $\omega_1(t)$ is replaced by

$$\omega_2(t) = \sup_{t_0 \leqslant t} \omega_1(t_0) ,$$ illustrated in Diagram (3) below,

the function obtained is no longer small, it being a well known fact that $\sum \frac{1}{p_n} = \infty$ (Niven and Zuckerman [9], Chapter 8).

Diagram 3



1.3.5 A function which, although below a small function, is not small

It is now possible to describe the function promised in Subsection 1.3.1, a function ω_2 such that $\omega_2(t) \leqslant \omega_1(t)$ for all t, ω_1 small, but ω_2 not small. Let ω_1 be as in Subsection 1.3.4 and define

$$\begin{cases} \omega_2\left(\frac{1}{p_n}\right) = \frac{1}{p_{n+1}} \\ \omega_2\ (t) = t^2 \quad \text{if} \quad t \neq \frac{1}{p_n} \quad \text{for any} \quad n\ . \end{cases}$$

Then $\omega_2(t) \leq \omega_1(t)$ for all t since $p_{n+1} \geq p_n + 1$.

But $\sum\omega_2^r(t)$ diverges for $t = \frac{1}{\sqrt[m]{p_n}}$; $m = 2^q$, $q \in \mathbb{N}$, $m \in \mathbb{N}$.

1.3.6 Composite small functions

Let T be an interval $\{t : 0 \leq t \leq t_0\}$, with ω_1 a small function on T . If φ is a positive increasing function defined on T such that for each $t \in T$ $\sigma_\varphi(t) = \sum \varphi(\omega^n(t)) < \infty$, then $\omega_2 = \varphi \circ \omega \circ \varphi^{-1}$ is a small function. To see this let

$$\sigma(t) = t + \omega_2(t) + \omega_2^2(t) + \ldots$$

$$= t + \sum\varphi[\omega^n(\varphi^{-1}(t))]$$

and note that then $\sigma(t) < \infty$ since $\sigma_\varphi(t) < \infty$.

1.4 Relative smallness of functions

The concept of a small function arose originally out of the need for a notion of relative smallness of functions. Pták [10] makes precise the idea of one function being "smaller than" another in the following way:

1.4.1 Definition Let t_0 be any positive number and denote by T the set $\{t : 0 < t < t_0\}$; denote by $S(T)$ the set of all functions q defined on T such that

$$q(t) > 0 \quad \text{for all} \quad t \in T \text{ , and} \quad \lim_{t \to 0} q(t) = 0 \text{ .}$$

Then, given two functions p and q in $S(T)$, p is said to be *smaller than* q if there exists a function ω defined on T with the following properties:

(i) $p(t) = q(\omega(t))$ for each $t \in T$;

(ii) the series $t + \omega(t) + \omega^2(t) + \ldots$ converges for each $t \in T$.

1.4.2 The particular case of linear functions

If $p(t) = \alpha t$ and $q(t) = \beta t$, then p is smaller than q if and only if $\alpha < \beta$.

Applications of this idea will be found in Section 6.

1.5 The cumulative error function and functional equation

Returning now to the definition of small function (1.1.1), let us denote by $\sigma(t)$ the sum of the series $t + \omega(t) + \omega^2(t) + \ldots$. We shall often refer to σ simply as the *cumulative error function* meaning, of course, the function corresponding to ω in this particular way. The function σ clearly satisfies the following functional equation:

$$\sigma(t) - t = \sigma(\omega(t)) \text{ .}$$

A consequence of this is the possibility of recovering ω if σ is given:

$$\sigma(\omega(t)) = \omega(t) + \omega^2(t) + \ldots$$

$$\text{whence} \quad \omega(t) = \sigma^{-1}[\omega(t) + \omega^2(t) + \ldots] \text{ ,}$$

$$\text{i.e.} \quad \omega(t) = \sigma^{-1}[\sigma(t) - t] \text{ .}$$

1.6 Formula for the partial sums of the cumulative error function

Although the knowledge of explicit formulas for ω and the corresponding σ already yields a considerable amount of information about the process as a whole, in applications it is necessary to have estimates for a finite number of steps. At first glance this seems to present difficulties since the iterated functions ω^n rapidly become fairly complicated. Nevertheless, in certain instances it is possible to obtain an explicit formula for the partial sums $t + \omega(t) + \dots + \omega^n(t)$. This also provides, of course, explicit expressions for $\omega^n(t)$. In order to give an example of the process involved we examine in detail the function $z \to \frac{\beta}{\gamma - z}$ where $\gamma^2 > 4\beta$, referred to in Subsection 1.2.3. This is interesting in that it leads to the small function defined by $\omega(t) = t\frac{\gamma + t - v}{\gamma - t + v}$, also mentioned in 1.2.3.

The rest of this section will be devoted to substantiating the above remarks and deriving the desired formula for the partial sums of the σ corresponding to $\omega(t) = t\frac{\gamma + t - v}{\gamma - t + v}$. For the sake of simplicity, instead of following Pták's approach [15] via the theory of continued fractions, which admittedly gives a clear indication of the evolution of the various results involved, we shall merely derive these results by means of ordinary mathematical induction.

1.6.1 Lemma Let β and γ be two positive numbers such that $\gamma^2 > 4\beta$. Suppose that z_0 is a positive number which does not belong to a countable set of exceptional values, E, to be described below. Then we have the following:

(1) it is possible to define a sequence z_k such that

$$z_k = \frac{\beta}{\gamma - z_{k-1}} \quad \text{for} \quad k = 1,2,\ldots \quad ,$$

the sequence z_k being expressed explicitly as

$$z_n = \lambda_2 \cdot \frac{1 - aq^{n-1}}{1 - aq^n}$$

where $\lambda_1 = \frac{1}{2}(\gamma + \sqrt{\gamma^2 - 4\beta})$,

$$\lambda_2 = \frac{1}{2}(\gamma - \sqrt{\gamma^2 - 4\beta}) ,$$

$$q = \frac{\lambda_2}{\lambda_1} \quad \text{and} \quad a = \frac{\lambda_2 - z_0}{\lambda_1 - z_0} \quad ;$$

(2) the exceptional set E consists of the numbers

$$e_n = \lambda_1 \cdot \frac{1 - q^{n+1}}{1 - q^n} \quad \text{for} \quad n = 1,2,\ldots \; ;$$

(3) all these numbers are different from each other and satisfy the inequalities

$$\lambda_2 < \lambda_1 < e_n < \gamma .$$

Proof

(1) The case $n = 1$: $z_1 = \lambda_2 . \frac{1 - a}{1 - aq}$

$$= \lambda_2 \cdot \frac{1 - \dfrac{\lambda_2 - z_0}{\lambda_1 - z_0}}{1 - \dfrac{\lambda_2 - z_0}{\lambda_1 - z_0} \cdot \dfrac{\lambda_2}{\lambda_1}} ,$$

from which it follows that

$$z_1 = \frac{\lambda_1\lambda_2}{\lambda_1 + \lambda_2 - z_0}$$

$$= \frac{\beta}{\gamma - z_0} ,$$

so the result is proved for $n = 1$. Now let us assume that the statement is true for $n = k - 1$ and prove that it is then true for $n = k$:

$$\gamma - z_{k-1} = \gamma - \lambda_2 . \frac{1 - aq^{k-2}}{1 - aq^{k-1}} ;$$

hence $$\frac{\beta}{\gamma - z_{k-1}} = \frac{\lambda_1\lambda_2}{\lambda_1 + \lambda_2 - \lambda_2 . \dfrac{1 - aq^{k-2}}{1 - aq^{k-1}}} .$$

The right hand side reduces to $\lambda_2 \dfrac{1 - aq^{k-1}}{1 - aq^k}$, i.e. to z_k . Thus, by mathematical induction, the result is true for all n .

(2) The exceptional values for z_0 are, of course, those for which $1 - aq^n = 0$, i.e. those for which $1 - \dfrac{q\lambda_1 - z_0}{\lambda_1 - z_0} . q^n = 0$, leading to $z_0 = \lambda_1 . \dfrac{1 - q^{n+1}}{1 - q^n}$.

(3) The inequality relationships involving λ_2, λ_1, e_n and γ are easily derivable:

(i) $\lambda_2 < \lambda_1$ by definition ;

(ii) $1 > \dfrac{\lambda_2}{\lambda_1} = q$

whence $1 - q^n < 1 - q^{n+1}$, giving

$$\lambda_1 < \lambda_1 . \frac{1 - q^{n+1}}{1 - q^n} ,$$

i.e. $\lambda_1 < e_n$;

(iii) $$e_n = \lambda_1 \cdot \frac{1 - \left(\frac{\lambda_2}{\lambda_1}\right)^{n+1}}{1 - \left(\frac{\lambda_2}{\lambda_1}\right)^n}$$

$$= \lambda_2 \cdot \frac{\lambda_1^{n+1} - \lambda_2^{n+1}}{\lambda_1^n - \lambda_2^n} .$$

Now we wish to prove $e_n < \gamma$, i.e. $e_n < \lambda_1 + \lambda_2$, so let us consider $(\lambda_1 + \lambda_2)(\lambda_1^n - \lambda_2^n)$. This expands to $\lambda_1^{n+1} - \lambda_1\lambda_2^n + \lambda_2\lambda_1^n - \lambda_2^{n+1}$ where $\lambda_2\lambda_1^n - \lambda_1\lambda_2^n > 0$ since $q\lambda_1^{n+1} - q^n\lambda_1^{n+1} > 0$ $(0 < q < 1 , q = \frac{\lambda_2}{\lambda_1})$, and the result is verified.

Next we examine some of the properties of the function established in 1.6.1 .

1.6.2 **Lemma** Let β and γ be two positive numbers such that $\gamma^2 > 4\beta$ and set $z_* = \frac{1}{2}(\gamma - \sqrt{\gamma^2 - 4\beta})$.

For $z < z_*$ the function

$$M(z) = \frac{\beta}{\gamma - z}$$

has the following properties:

(i) $z < M(z) < z_*$;

(ii) if $t > 0$ and if $z < z_*$ is such that $M(z) - z = t$, then

$$M(z + t) - (z + t) = t \cdot \frac{\gamma + t - \sqrt{(\gamma + t)^2 - 4\beta}}{\gamma - t + \sqrt{(\gamma + t)^2 - 4\beta}} .$$

Proof (i) $z_*^2 - \gamma z_* = -\beta$, and since

$$M(z) - z = \frac{z^2 - \gamma z + \beta}{\gamma - z} ,$$ we obtain

$$M(z) - z = \frac{z^2 - \gamma z + \beta - (z_*^2 - \gamma z_* + \beta)}{\gamma - z}$$

$$= \frac{(z - z_*)(z + z_* - \gamma)}{\gamma - z} .$$

Now, for $z - z_* < 0$, we have

$$z + z_* - \gamma < 2z_* - \gamma < 0 , \text{ giving}$$

$$M(z) - z > 0 .$$

Similarly, $z_* - M(z) = \frac{1}{\gamma - z}[z_*\gamma - zz_* - \beta + (z_*^2 - \gamma z_* + \beta)]$

$$= z_* . \frac{z_* - z}{\gamma - z} ,$$

and $\frac{z_*}{\gamma - z} > \frac{z}{\gamma - z} = \frac{z . M(z)}{\beta} > \frac{z^2}{\beta}$,

so $z_* - M(z) > 0$.

(ii) Suppose $z < z_*$ and $M(z) - z = t > 0$. Write z' for $z + t = M(z)$ and observe that $\beta = z'(\gamma - z)$. It follows that

$$M(z') - z' = \frac{\beta - z'(\gamma - z')}{\gamma - z'}$$

$$= \frac{z't}{\gamma - z'}$$

$$= t . \frac{z + t}{\gamma - t - z} .$$

To eliminate z we observe that $\beta - (z + t)(\gamma - z) = 0$ or $z^2 - (\gamma - t)z + \beta - \gamma t = 0$, from which it is clear that

$$z = \tfrac{1}{2}[\gamma - t \pm \sqrt{(\gamma + t)^2 - 4\beta}] ,$$

and since the '+' possibility yields $z > z_*$, we conclude that

$$z = \tfrac{1}{2}[\gamma - t - \sqrt{(\gamma + t)^2 - 4\beta}] .$$

Replacing $\sqrt{(\gamma + t)^2 - 4\beta}$ by v , we write

$$z = \tfrac{1}{2}(\gamma - t - v) , \text{ so that}$$

$$M(z') - z' = t \cdot \frac{\gamma + t - v}{\gamma - t + v} \,.$$

We are now ready to compute the formula for the partial sums.

1.6.3 Theorem Let β and γ be two positive numbers such that $\gamma^2 > 4\beta$. For each $t > 0$ let $v = \sqrt{(\gamma + t)^2 - 4\beta}$ writing v_0 for $\sqrt{\gamma^2 - 4\beta}$, and set

$$\omega(t) = t \cdot \frac{\gamma + t - v}{\gamma - t + v} \,.$$

Then ω is a rate of convergence on the whole positive axis and possesses the following properties:

(1) For each natural number n and each $t > 0$ we have

$$t + \omega(t) + \ldots + \omega^n(t) = M^{n+1}(z_0) - z_0 \,,$$

where

$$M^{n+1}(z_0) = \frac{1}{2}(\gamma - v_0) \cdot \frac{1 - \frac{t + v - v_0}{t + v + v_0}\left(\frac{\gamma - v_0}{\gamma + v_0}\right)^n}{1 - \frac{t + v - v_0}{t + v + v_0}\left(\frac{\gamma - v_0}{\gamma + v_0}\right)^{n+1}} - \frac{1}{2}(\gamma - t - v)$$

and $z_0 = \frac{1}{2}(\gamma - t - v)$;

(2) The infinite sum is given by

$$\sigma(t) = \frac{1}{2}(t + v - v_0) \,.$$

Proof (1) Suppose that z_0, $0 < z_0 < z*$, is such that $M(z_0) - z_0 = t$. Then, according to 1.6.2, $M^2(z_0) - M(z_0) = \omega(t)$ and $M^k(z_0) - M^{k-1}(z_0) = \omega^{k-1}(t)$ for all $k \geqslant 1$.

It follows that

$$\begin{aligned} & t + \omega(t) + \omega^2(t) + \dots + \omega^n(t) \\ &= M(z_0) - z_0 + M^2(z_0) - M(z_0) + M^3(z_0) - M^2(z_0) + \dots \\ &\qquad \dots + M^{n+1}(z_0) - M^n(z_0) \\ &= M^{n+1}(z_0) - z_0 \,. \end{aligned}$$

Now $z_0 = \frac{1}{2}(\gamma - t - v)$ as for z in 1.6.2 and $M^{n+1}(z_0)$ is, of course, the n-th iterate of the continued fraction $z_k = \dfrac{\beta}{\gamma - z_{k-1}}$ defined in 1.6.1.

This means that

$$M^{n+1}(z_0) = z_n = \lambda_2 \,.\, \frac{1 - aq^n}{1 - aq^{n-1}} \,.$$

But $\lambda_2 = \frac{1}{2}(\gamma - v_0)$ and $a = \dfrac{\lambda_2 - z_0}{\lambda_1 - z_0}$

$$= \frac{\frac{\gamma - v_0}{2} - z_0}{\frac{\gamma + v_0}{2} - z_0}$$

$$= \frac{t + v - v_0}{t + v + v_0} \,,$$

so the partial sums have the desired form.

(2) For the partial sums developed above write

$$\sigma_n'(t) = \frac{1 - aq^n}{1 - aq^{n+1}} \,,$$

then $\displaystyle\lim_{n\to\infty} \sigma_n'(t) = \frac{1}{q}$

$$= \frac{\gamma - v_0}{\gamma + v_0} \,.$$

Hence

$$\sigma(t) = \frac{1}{2}(\gamma - v_0) \,.\, \frac{\gamma + v_0}{\gamma - v_0} - \frac{1}{2}(\gamma - t - v)$$

$$= \frac{1}{2}(t + v - v_0) \,.$$

Section 2

THE INDUCTION THEOREM

The Induction Theorem evolved from Pták's re-moulding of the approximation technique basic to the Closed Graph Theorem [10]. It works, however, under much weaker hypotheses, namely those of the metric generalization of the Closed Graph Theorem. To explain what we mean by the metric generalization of the Closed Graph Theorem, let us consider this theorem in its open mapping form and in the special case of normed spaces.

2.1 The metric generalization of the Closed Graph Theorem

Let E be a normed space and M a linear mapping onto another normed space F. Consider the unit cell U of E and its image MU. If F possesses the Baire property, the set MU cannot be nowhere dense in F, so that the closure $(MU)^-$ is a neighbourhood of zero in F. This observation is the starting point of an iteration process which proves for E Banach and M closed, the inclusion

$$(1 + \varepsilon)(MU) \supset (MU)^- ,$$

which in turn shows that the mapping M is open.

In fact, the inclusion $(MU)^- \supset V$, a neighbourhood of zero in F, is more than we need in order to set up a convergent iterative process. $(MU)^- \supset V$ means that points of set V may be arbitrarily well approximated by points of

MU . The degree of approximation, however, need not be so good; it suffices to approximate points of V by those of MU to within αV where $0 < \alpha < 1$, and this is the essence of the Induction Theorem.

The above hypothesis concerning the degree of approximation has a linear form and, of course, in order to achieve greater generality, we should try to make it non-linear. This requires the notion of relative smallness of functions dealt with in Subsection 1.4. Clearly, the assumption $\alpha < 1$ in the generalized Open Mapping Theorem above simply means that αt is smaller than the identity function t .

2.2 Uniform Lower Semicontinuity

Pták's initial formulation of the Induction Theorem [10] is based on the above ideas as well as the concept of lower semicontinuity to be defined below. Firstly, let us note that the following notation for spherical neighbourhoods is to be used here and in the sequel:

if (E,d) is a metric space, then

$B(x,t) = \{y \in E : d(y,x) < t\}$ and, for $X \subset E$,

$B(X,t) = \{y \in E : d(y,x) < t$ for some $x \in X\}$.

2.2.1 Definition Let (E,d) be a metric space. Suppose we are given, for each sufficiently small positive t , a set $Z(t) \subset E$. We shall say that the system $Z(\cdot)$ is *uniformly lower semicontinuous* if there exist two functions $p, q \in S$ (S as in Definition 1.4.1), the function p being small with respect to q, such that for each sufficiently small t and each $x \in Z(q(t))$ the intersection $B(x,t) \cap Z(p(t))$ is non-void, i.e. we have the inclusion $Z(q(t)) \subset B\big(Z(p(t)),t\big)$.

For the sake of simplicity let us talk merely of *semicontinuity*, taking this always to mean *uniform lower semicontinuity* as defined above.

Given a family $Z(\cdot)$ as above, one further prerequisite for the theorem is the definition of the limit of this family.

2.2.2 Definition The *limit of the family* $Z(\cdot)$ is given by

$$Z(0) = \bigcap_{s>0} \Big(\bigcup_{t \leq s} Z(t) \Big)^{-} .$$

Often $Z(t)$ is monotone and then $Z(0) = \bigcap_{t>0} \overline{Z(t)}$.

2.3 The Induction Theorem in its different forms

2.3.1 Theorem Let $Z(\cdot)$ be a semicontinuous system of approximate sets in a complete metric space E with distance function d . Then, for each sufficiently small $t > 0$,

$$Z(q(t)) \subset B(Z(0), \sigma(t)) .$$

Proof Let $x \in Z(q(t))$. Then, by hypothesis, there exists an $x_1 \in B(x,t) \cap Z(p(t))$. Since $p(t) = q(\omega(t))$ as in Definition 1.4.1, we have $x_1 \in Z\big(q(\omega(t))\big)$, so that there exists an $x_2 \in B(x_1, \omega(t)) \cap Z\Big(p(\omega(t))\Big)$.

Now $p(\omega(t)) = q(\omega^2(t))$ so there must exist an $x_3 \in B(x_2, \omega^2(t)) \cap Z\Big(p(\omega^2(t))\Big)$. Proceeding by induction we obtain a sequence

$$x_{n+1} \in B(x_n, \omega^n(t)) \cap Z\Big(p(\omega^n(t))\Big) .$$

The distances $d(x_{n+1}, x_n)$ being dominated by the terms of the convergent series $t + \omega(t) + \omega^2(t) + \ldots$ and E being complete, the sequence $\{x_n\}$ converges to an element $x_\infty \in E$. Clearly $x_\infty \in Z(0)$ and, furthermore,

$$\begin{aligned} d(x,x_\infty) &\leqslant d(x,x_1) + d(x_1,x_2) + \ldots \\ &< t + \omega(t) + \ldots , \end{aligned}$$

i.e. $d(x,x_\infty) < \sigma(t)$.

The statement of the theorem can in fact be streamlined, using rates of convergence explicitly, by putting $p = \omega$ and $q(t) = t$. The proof remains much the same (Ptak [11]).

2.3.2 <u>Theorem</u> Let (E,d) be a complete metric space, T an interval $\{t : 0 < t < t_0\}$ and ω a rate of convergence on T . For each $t \in T$ let $Z(t)$ be a subset of E . Then

$$Z(t) \subset B\Big(Z(\omega(t)),t\Big) \text{ for each } t \in T \Rightarrow Z(t) \subset B(Z(0), \sigma(t)) \text{ for each } t \in T.$$

There is another form of this theorem which is simpler formally but not as convenient in applications (Pták [12]). In it (E,d) is a metric space with d defined in the following way:

if $X, Y \subset E$ then $d(X,Y) = \inf\{r : X \subset B(Y,r)\}$

(a distance which is not symmetric and may be infinite).

2.3.3 **Theorem** If (E,d) is complete then

$$d\Big(Z(t),\ Z(\omega(t))\Big) \leqslant t \ \text{ for small } \ t \Rightarrow d(Z(t),\ Z(0)) \leqslant \sigma(t)\ .$$

Finally there is a form in which Pták uses the Induction Theorem in his proof of Moser's Theorem [16]. Mention of this will be made in Section 5.

2.3.4 **Theorem** Given E a complete metric space, T an interval of the form $\{t : 0 < t < t_0\}$ and ω a rate of convergence on T. Let φ be a positive increasing function defined on T such that, for each $t \in T$, $\sigma_\varphi(t) = \sum \varphi(\omega^n(t)) < \infty$. If $W(\cdot)$ is a family of subsets of E such that $W(t) \subset B\Big(W(\omega(t)),\ \varphi(t)\Big)$ for each $t \in T$, then $W(t) \subset B(W(0),\ \sigma_\varphi(t))$ for each $t \in T$.

Proof Set $Z(t) = W(\varphi^{-1}(t))$

so that $Z(\omega(t)) = W\Big(\omega(\varphi^{-1}(t))\Big)$ and $W(t) = Z(\varphi(t))$.

$$\text{Then} \quad W(t) \subset B\Big(W(0),\ \sigma_\varphi(t)\Big)$$

$$\Longleftrightarrow Z(\varphi(t)) \subset B(Z(0),\ \sigma_\varphi(t))\ .$$

But we are given $W(t) \subset B\Big(W(\omega(t)),\ \varphi(t)\Big)$,

$$\text{i.e.} \quad Z(\varphi(t)) \subset B\Big(Z\big(\omega(\varphi(t))\big),\ \varphi(t)\Big)\ .$$

So, by applying the Induction Theorem (2.3.1), we have that

$$Z(\varphi(t)) \subset B\Big(Z(0), \sigma(\varphi(t))\Big)\ ,$$

$$\text{i.e.} \quad W(t) \subset B(W(0),\ \sigma_\varphi(t))\ .$$

In concluding this section on the Induction Theorem, let us examine the way in which this theorem is to be applied (Pták [12]).

2.4 Principles of application of the Induction Theorem

To see how the Induction Theorem is used we refer back to the problem posed in the introduction concerning the construction of a point x belonging to a given set Z.

Supposing we have an approximation x of order s and are allowed to move from x to a distance not greater than r, can we find, within $B(x,r)$, an approximation of a much better order s', "much better" meaning that $s' = \omega(s)$? If so, then the family of sets $Z(\cdot)$, where $Z(s)$ is the set of approximations of order s or better, satisfies the hypothesis of the Induction Theorem. We have thus $Z(t) \subset B(Z(0), \sigma(t))$ for sufficiently small t. Hence $Z(0)$ is non-empty if at least one $Z(t)$ is non-empty. In this way the Induction Theorem makes it possible to reduce the analysis involved in existence proofs to just the verification that it is possible to pass from a given approximation to a much better one by choosing a suitable element within a given distance.

Suppose we are to find a solution to an equation $f(u) = 0$. Given a positive function m which measures how close $f(x)$ is to zero, it is natural to define the family $Z(\cdot)$ in such a way that

$$Z(t) \subset \{x \in E : m(f(x)) \leqslant \varphi(t)\} ,$$

where φ is a positive function which tends to zero with t.

Application of the Induction Theorem consists essentially of the two steps described below.

(1) Given a fixed $u \in E$ and a positive number t, find or estimate

$$\inf \{m(f(u')) : u' \in B(u,t)\} .$$

Compare this infimum (or its estimate) with the value $m(f(u))$. In favourable circumstances we can show that the above infimum is small as compared with $m(f(u))$. This means that if we have the estimate

$$m(f(u)) \leqslant p \Rightarrow \inf \{m(f(u')) : u' \in B(u,t)\} \leqslant h(p,t) ,$$

where h is a positive function, then we can find a small function ω such that φ and ω satisfy

$$h(\varphi(t),t) \leqslant \varphi(\omega(t)) .$$

We may then assert that given u with $m(f(u)) \leqslant \varphi(t)$, there exists, within a distance t, a point u' for which

$$\text{(A)} \qquad m(f(u')) \leqslant \varphi(\omega(t)) .$$

(2) Having fixed ω, construct the cumulative error function

$$\sigma(t) = t + \omega(t) + \omega^2(t) + \dots .$$

Now suppose that $d(u,u_0) \leqslant \alpha - \sigma(t)$, then using the functional equation in 1.5, for each $u' \in B(u,t)$, we have

$$\text{(B)} \qquad d(u',u_0) \leqslant d(u,u_0) + d(u',u) \leqslant \alpha - \sigma(t) + t ,$$

$$\text{i.e. } d(u',u_0) \leqslant \alpha - \sigma(\omega(t)) .$$

Here α is subject to the obvious requirement that $\alpha - \sigma(t) > 0$.

(A) and (B) together imply the inclusion

$$Z(t) \subset B\Big(Z(\omega(t)),\ t\Big)$$

if $Z(t)$ is defined as follows:

$$Z(t) = \{x \in E : m(f(x)) \leq \varphi(t),\ d(x,u_0) \leq \alpha - \sigma(t)\} .$$

The value of α is eventually determined by the requirement that at least one $Z(t)$ be non-void.

The above process will be amply illustrated in Sections 5 and 6.

Section 3

RELATIONSHIPS BETWEEN THE INDUCTION THEOREM AND SOME CLASSICAL THEOREMS OF ANALYSIS

3.1 The Closed Graph Theorem

The most vital aspect of the Induction Theorem lies in the fact that it represents an abstract form of an approximation process very useful in analysis. In this respect it plays the same role as the Closed Graph Theorem which is nothing more than the abstract description of an approximation process, the theorem itself saving us the inconvenience of going through this process in concrete situations. The Induction Theorem is in fact stronger than the Closed Graph Theorem. Indeed, the Closed Graph Theorem can, in a sense, be seen as a limiting case of the Induction Theorem for an infinitely fast rate of convergence. Pták [12] proves a strengthening of the Closed Graph Theorem, formulated in such a way that it becomes an immediate consequence of the Induction Theorem. Before being able to substantiate this last remark we require a few basic results concerning closed relations.

3.1.1 **Closed Relations** A relation R from a set E into a set F is a subset of $E \times F$. We write $y = Rx$ and $x \in R^{-1}y$ if $(x,y) \in R$. Similarly, if $C \subset F$ then $R^{-1}C$ is the set of all $x \in E$ such that $(x,c) \in R$ for some $c \in C$. In other words,

$$R^{-1}C = \bigcup\{R^{-1}c : c \in C\} = \{x \in E : Rx \cap C \neq \emptyset\} .$$

For $A \subset E$, $RA = \{Ra : a \in A\} = \{y \in F : R^{-1}y \cap A \neq \emptyset\}$.

It is often more useful to restate inclusions concerning R in terms of R^{-1} and vice-versa. In particular, we have the following result:

3.1.1.1 **Lemma** The conditions

$$x \in B(R^{-1}y,r) , \quad y \in RB(x,r) , \quad R^{-1}y \cap B(x,r) \neq \emptyset$$

are equivalent.

Restated for R^{-1} these conditions are altered as follows:

3.1.1.2 **Lemma** The conditions

$$y \in B(Rx,r) , \quad x \in R^{-1}B(y,r) , \quad Rx \cap B(y,r) \neq \emptyset$$

are equivalent.

Further requirements before we can proceed to the theorem are definitions of the notions of uniform almost openness and uniform openness.

3.1.1.3 **Definition** Let E and F be metric spaces with R a closed subset of $E \times F$. Let $D(R)$ be the domain of R . Then R is said to be *uniformly almost open* if for each $r > 0$ there exists a positive number $q(r)$ such that

$$(RB(x,r))^{-} \supset B(Rx, q(r)) \quad \text{for each} \quad x \in D(R) .$$

3.1.1.4 **Definition** Given E, F and R as in 3.1.1.3 R is said to be *uniformly open* if for each $r > 0$ there exists a positive number $q(r)$ such that, for each $r' > R$,

$$RB(x,r') \supset B(Rx, q(r)) \text{ for each } x \in D(R) .$$

Now we are ready to prove the Closed Graph Theorem by means of the Induction Theorem.

3.1.2 **Theorem** Let E be a complete metric space and F any metric space. Let R be a closed subset of $E \times F$. If the relation R is uniformly almost open, then it is uniformly open.

Proof Let $r > 0$, $r' > r$ and $x \in D(R)$ be fixed and consider an arbitrary $y_0 \in B(Rx, q(r))$. For each $t > 0$ set

$$Z(t) = R^{-1}B(y_0,t) .$$

Note that, by Lemma 3.1.1.2,

$$Z(t) = \{x \in E : Rx \cap B(y_0,t) \neq \emptyset\}$$

or $$Z(t) = \{x \in E : y_0 \in B(Rx,t)\} .$$

Since R is closed it follows from Definition 2.2.2 that

$$Z(0) = R^{-1}y_0 .$$

In order to prove the theorem we must now verify the inclusion

$$(1) \quad Z(q(r)) \subset B(Z(0), r') ,$$

because, if $x \in Z(q(r))$, (1) implies $x \in B(Z(0), r')$, i.e. $x \in B(R^{-1}y_0, r')$ whence $y_0 \in RB(x,r')$ by 3.1.1.1.

Since y_0 was an arbitrary point in $B(Rx,\ q(r))$ we would then have $B(Rx,\ q(r)) \subset RB(x,r')$.

To prove (1) take $t > 0$, $s > 0$, and note that if $z \in Z(q(t))$, then by 3.1.1.2 $y_0 \in B(Rz,\ q(t)) \subset (RB\ (z,t))^{-}$, so that $B(y_0,s) \cap RB(z,t) \neq \emptyset$. Now, if $w \in B(z,t)$ is such that $B(y_0,s) \cap Rw \neq \emptyset$, then $w \in Z(s)$ and $z \in B(w,t) \subset B(Z(s),\ t)$, leading to the fact that

$$(2) \quad Z(q(t)) \qquad B(Z(s),\ t).$$

Now we introduce the function $q^*(t) = \min(q(t),\ r^{-1}q(r)t)$, so that $q^*(r) = q(r)$ and $q^*(t) \to 0$. Since $q^*(t) \leqslant q(t)$, we have for each $t > 0$, $s > 0$,

$$(3) \quad Z(q^*(t)) \subset B(Z(s),\ t) \quad \text{by (2)} .$$

Now choose $0 < \varepsilon < 1$ so as to have $\frac{r}{1-\varepsilon} = r'$. Then, setting $s = q^*(\varepsilon t)$ in (3), we obtain

$$Z(q^*(t)) \subset B\Big(Z(q^*(\varepsilon t)),\ t\Big) \quad \text{for each} \quad t > 0 .$$

It follows from the Induction Theorem (2.3.1) that

$$Z(q^*(t)) \subset B\Big(Z(0),\ \frac{1}{1-\varepsilon} \cdot t\Big) = B\Big(R^{-1}y_0,\ \frac{1}{1-\varepsilon} \cdot t\Big)$$

for each $t > 0$.

In particular, for $t = r$, we obtain

$$Z(q(r)) \subset B(R^{-1}y_0,\ r') = B(Z(0),\ r') ,$$

which completes the proof.

3.2 The Banach Fixed Point Theorem

Pták also shows in his paper [12] that the Induction

Theorem is very closely related to the well known Banach Fixed Point Theorem. In fact, the Induction Theorem provides a slight generalization of Banach's theorem.

3.2.1 **Theorem** Let (E,d) be a complete metric space and f a mapping of E into itself such that, for each x, y in E, $d(f(x), f(y)) \leqslant kd(x,y)$ where k is a fixed number, $0 < k < 1$. Then there exists a unique $\xi \in E$ such that $f(\xi) = \xi$.

Proof For each $t > 0$ set

$$Z(t) = \{x : d(x, f(x)) < t\} .$$

Then, of course, $Z(0) = \{x : x = f(x)\}$.

Now it is easy to see that if $x \in Z(t)$ and $x' = f(x)$, so that $d(x,x') < t$, then $Z(t) \subset B(Z(kt), t)$, since

$$d(x', f(x')) = d(f(x), f(x')) \leqslant kd(x,x') = kd(x, f(x)) < kt .$$

So the Induction Theorem (2.3.2) applies with $\omega(t) = kt$, giving $Z(0) \neq \emptyset$. To prove that $Z(0)$ is a singleton is a simple matter: if ξ and $\xi' \in Z(0)$ then

$$\begin{aligned} d(\xi,\xi') &= d(f(\xi), f(\xi')) \\ &\leqslant kd(\xi,\xi') \\ &< d(\xi,\xi') , \end{aligned}$$

which is impossible unless $d(\xi,\xi') = 0$.

We now turn to another basic result of contractive type which can be proved via the Induction Theorem, *viz.* Cantor's Intersection Theorem.

3.3 The Cantor Intersection Theorem

The Induction Theorem is used to provide two proofs of the Cantor Intersection Theorem (Simmons [17], Chapter 2).

3.3.1 **Theorem** Let E be a complete metric space with $\{F_n\}$ a decreasing sequence of non-empty closed subsets of E such that $d(F_n)$ (the diameter of F_n) $\to 0$.

Then $F = \bigcap_1^\infty F_n$ contains exactly one point.

Proof Let $T = \{t : 0 < t < d(F_1)\}$ and define a small function ω on T by $\omega(t) = \alpha t$, $0 < \alpha < 1$. For each $t \in T$ let $Z(t) = F_{n(t)}$ where $n(t)$ = the smallest n such that $d(F_n) \leqslant t$. (This definition is meaningful since $d(F_n) \to 0$.) Now $\sigma(t)$ converges and $Z(t) \subset B\Big(Z(\omega(t)), t\Big)$ so the Induction Theorem applies giving us $Z(0) \neq \emptyset$, i.e. $\bigcap_1^\infty F_n \neq \emptyset$ and, since $d(F_n) \to 0$, $\bigcap_1^\infty F_n$ is a singleton.

3.3.2 **Alternative proof** Pick a subsequence of $\{F_n\}$, denoted again by $\{F_n\}$, such that $\sum d(F_n) < \infty$ and $d(F_n) \searrow 0$. Define $Z(t) = F_{n(t)}$ as above and define $\omega(t) = d(F_{n(t)})$. Then, as before, $\sigma(t)$ converges and $Z(t) \subset B\Big(Z(\omega(t)), t\Big)$, etc.

Here we have used the fact, mentioned in Subsection 1.3.3, that if $\sum a_n < \infty$, $a_n \searrow 0$, then by defining $\omega(t)$ = the first a_n such that $a_n < t$, we generate a small function.

One further result of the above type is Baire's Category

Theorem.

3.4 The Baire Category Theorem

The Induction Theorem can be used to modify the usual proof of the Baire Category Theorem (Kelley [7], Chapter 6).

3.4.1 Theorem Let E be a complete metric space. Then the intersection of any countable family of open dense subsets of E is itself dense in E, i.e. E is a Baire space .

Proof Let $\{U_n\}$ be a sequence of open, dense subsets of E and G an arbitrary open, non-empty subset of E . It must be shown that $G \cap \bigcap_1^\infty U_n \neq \emptyset$.

Choose inductively an open set V_1 such that $\overline{V}_1 \subset G \cap U_1$ and $d(V_1) < 1$. Then, for each $n > 1$, choose V_n such that $\overline{V_n} \subset V_{n-1} \cap U_n$ and $d(V_n) < \frac{1}{n}$. This choice is possible because U_n is dense and open.

Let $T = \{t : 0 < t < 1\}$ and define ω , a small function on T , by $\omega(t) = \frac{1}{2}t$. Denote by $\left[\frac{1}{t}\right]$ the integral part of $\frac{1}{t}$ and let $n(t) = \left[\frac{1}{t}\right] + 1$.

For each t let $Z(t) = V_{n(t)}$, i.e. $Z(t) = V_{\left[\frac{1}{t}\right]+1}$.

Note that $Z(\omega(t)) = V_{\left[\frac{2}{t}\right]+1}$ and, since $\{V_n\}$ is a contracting sequence, we have $V_{\left[\frac{2}{t}\right]+1} \subseteq V_{\left[\frac{1}{t}\right]+1}$.

But $d\left(V_{\left[\frac{1}{t}\right]+1}\right) < \dfrac{1}{\left[\frac{1}{t}\right] + 1} \leqslant t$, so we have the result that

$$V_{\left[\frac{1}{t}\right]+1} \subset B\left(V_{\left[\frac{2}{t}\right]+1} , t\right) ,$$

i.e. $$Z(t) \subset B\Big(Z(\omega(t)), t\Big) .$$

Now, by applying the Induction Theorem (2.3.2), we obtain $Z(0) \neq \emptyset$, i.e. $\bigcap_1^\infty V_n \neq \emptyset$. And, since $\overline{V_{n+1}} \subset G \cap U_n$, it follows that

$$G \cap \bigcap_1^\infty U_n \neq \emptyset .$$

Section 4

APPLICATION OF THE INDUCTION THEOREM TO PROVIDE EXTENDED CONTRACTION PRINCIPLES

4.1 A small function generalization of Banach's Fixed Point Theorem

The Banach Fixed Point Theorem was generalized by means of the Induction Theorem in Subsection 3.2. Replacing the constant k by the function $\omega(d(\cdot))$, ω a rate of convergence, we can achieve an even greater degree of generalization. This is an idea which Pták himself does not appear to have considered.

4.1.1 **Theorem** Let (E,d) be a complete metric space and f a mapping of E into itself such that, for all $x,y \in E$, $d(f(x), f(y)) \leq \omega(d(x,y))$, ω a monotone increasing rate of convergence.
Then there exists a unique $\xi \in E$ such that $f(\xi) = \xi$.

Proof For each $t > 0$ set

$$Z(t) = \{x : d(x, f(x)) < t\} .$$

Then, as in 3.2.1, $Z(0) = \{x : x = f(x)\}$.
And this time it will suffice to show that $Z(t) \subset B\Big(Z(\omega(t)), t\Big)$.

Again, if $x \in Z(t)$ set $x' = f(x)$ so that $d(x,x') < t$. And showing that $x' \in Z(\omega(t))$ is straightforward since, because ω is monotone,

$$d(x', f(x')) = d(f(x), f(x')) \leq \omega(d(x,x')) \leq \omega(t) .$$

So $Z(t) \subset B\Big(Z(\omega(t)), t\Big)$ and the Induction Theorem applies. Once again the uniqueness of ξ is easy to prove. We do, however, require an approach different from that at the end of 3.2.1. If ξ and ξ' are fixed points then, for all $n \in \mathbb{N}$,

$$d(\xi,\xi') = d(f(\xi), f(\xi')) = d(f^n(\xi), f^n(\xi')) .$$

$$\text{Hence} \quad d(\xi,\xi') \leqslant \omega\Big(d(f^{n-1}(\xi), f^{n-1}(\xi'))\Big)$$
$$\leqslant \omega^n(d(\xi,\xi'))$$
$$\xrightarrow[\infty]{n} 0 \quad \text{since } \omega \text{ small.}$$

In the same vein, the Induction Theorem also enables us to generalize some work on contraction done by Edelstein [4].

4.2 Edelstein's Extended Contraction Principle

4.2.1 **Definition** Given (E,d) a metric space, $f : E \to E$ is said to be *locally contractive* if, for every $x \in E$, there exists ε, λ $(\varepsilon > 0\ ,\ 0 \leqslant \lambda < 1)$ such that $p,q \in B(x,\varepsilon) \Rightarrow d(f(p), f(q)) < \lambda d(p,q)$.

f is called (ε,λ)-*uniformly locally contractive* if it is locally contractive and both ε and λ are independent of x .

4.2.2 **Definition** A metric space (E,d) is said to be ε-chainable if, for every $a,b \in E$, there exists an ε-chain, i.e. a finite set of points $a = x_0, x_1, \ldots, x_n = b$, such that $d(x_{i-1}, x_i) < \varepsilon$ $(i = 1,2,\ldots,n)$.

Edelstein's Theorem is stated without proof.

4.2.3 **Theorem** Let (E,d) be a complete metric ε-chainable space, with f a mapping of E into itself which is (ε,λ)-uniformly locally contractive. Then there exists a unique point $\xi \in E$ such that $f(\xi) = \xi$.

Let us show how this result can be extended in a fairly natural way if λ is replaced by ω, a Pták rate of convergence.

4.2.4 **Definition** Given (E,d) a metric space, $f : E \to E$ will be called (ε,ω)-*uniformly locally contractive* if there exists ε,ω ($\varepsilon > 0$, ω small, monotone increasing on $(0,2\varepsilon)$) such that, for every $x \in E$,

$$p,\ q \in B(x,\varepsilon) \Rightarrow d(f(p),\ f(q)) \leqslant \omega(d(p,q)) .$$

4.2.5 **Theorem** Let E be a complete metric ε-chainable space, with f a mapping of E into itself which is (ε,ω)-uniformly locally contractive. Then there exists a unique point $\xi \in E$ such that $f(\xi) = \xi$.

Proof Let x be an arbitrary point of E and consider the ε-chain $x = x_0,\ x_1,\ \ldots,\ x_n = f(x)$. By the Triangle Inequality $d(x,\ f(x)) \leqslant \sum_1^n d(x_{i-1},\ x_i) < n\varepsilon$.

Also, by uniform local contractivity,

$$d(f(x_{i-1}),\ f(x_i)) \leqslant \omega(d(x_{i-1},x_i)) \leqslant \omega(\varepsilon)$$

and, by induction,

$$d(f^m(x_{i-1}),\ f^m(x_i)) \leqslant \omega\Big(d(f^{m-1}(x_{i-1}),\ f^{m-1}(x_i))\Big),$$

i.e. $d(f^m(x_{i-1}), f^m(x_i)) \leqslant \omega^m(\varepsilon)$.

From this ,

$$d(f^m(x), f^{m+1}(x)) \leqslant \sum_1^n d(f^m(x_{i-1}), f^m(x_i)) \leqslant n\omega^m(\varepsilon) .$$

It follows that $\{f^i(x)\}$ is Cauchy:

$$0 < j < k \Rightarrow d\left(f^j(x), f^k(x)\right) \leqslant \sum_{i=j}^{k-1} d(f^i(x), f^{i+1}(x))$$

$$\leqslant n.[\omega^j(\varepsilon) + \dots + \omega^{k-1}(\varepsilon)]$$

$$\xrightarrow[\infty]{j} 0 ,$$

since if $s_n = \sum_1^n \omega^r(\varepsilon)$ then $s_n < \infty$ (ω small) and this implies $s_{k-1} - s_j \to 0$.

So by completeness of E we know that $\xi = \lim_{i\to\infty} f^i(x)$ exists, and from the continuity of f it follows that

$$f(\lim_{i\to\infty} f^i(x)) = \lim_{i\to\infty} f(f^i(x)) = \lim_{i\to\infty} f^{i+1}(x) = \lim_{i\to\infty} f^i(x) .$$

Now all we need do to complete the proof is to set $\xi = \lim_{i\to\infty} f^i(x)$ and show that there is no other point ξ' satisfying $f(\xi') = \xi'$. Suppose such a point $\xi' \neq \xi$ exists and let $\xi = x_0, x_1, \dots, x_p = \xi'$ be an ε-chain.

Then $d(f(\xi), f(\xi')) = d(f^q(\xi), f^q(\xi'))$ (for all $q \in \mathbb{N}$)

$$\leqslant \sum_{i=1}^{p} d(f^q(x_{i-1}), f^q(x_i)) \text{ as in (2) .}$$

Thus $d(f(\xi), f(\xi')) \leqslant p\omega^q(\varepsilon) \xrightarrow[\infty]{q} 0$, which is impossible.

Hence $\xi = \xi'$ and the result is proved.

Section 5

THE INDUCTION THEOREM APPLIED TO NEWTON'S METHOD

In this section we shall employ the Induction Theorem to establish a small function which gives the rate of convergence of Newton's Method. We shall then derive a formula for the partial sums of the corresponding cumulative error function. Finally, we shall mention Pták's use of the Induction Theorem to improve Moser's Theorem concerning a modification of Newton's Method. The approach followed will be that of Pták in his papers [13] and [14], with alterations in notation and format and more detailed explanations wherever greater clarity could be achieved.

5.1 Newton's Process in Banach Spaces

We begin by sketching Newton's Method for functions $f : \mathbb{R} \to \mathbb{R}$. Suppose that f has first and second order derivatives on the interval $a \leqslant x \leqslant b$, that $f(a)$ and $f(b)$ are of opposite sign, and that $f'(x)$ and $f''(x)$ are each of constant sign on $a \leqslant x \leqslant b$. Then there is a unique x between a and b for which $f(x) = 0$. If x_n $(n = 1,2,\ldots)$ is an approximation to the solution, then

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

is a better approximation.

Now let us extend the above ideas to Banach spaces.

If E and F are normed spaces and $f : E \to F$, to solve $f(x) = 0$ requires the notion of $f'(x)$ regarded as a linear operator.

5.1.1 **Definition** Let E and F be Banach spaces (both real or complex), U an open subset of E, and f and g two mappings of U into F. We say that f and g are *tangent* at a point $y \in U$ if

$$\lim_{x \to y,\ x \neq y} \frac{\| f(x) - g(x) \|}{\| x - y \|} = 0 ;$$

this implies of course that $f(y) = g(y)$.

5.1.2 **Definition** We say that a continuous mapping f of U into F is *differentiable* at the point $y \in U$ if there is a linear mapping h of E into F such that $x \to f(y) + h(x - y)$ is tangent to f at y. This mapping h is unique; it is called the derivative of f at the point y and written $f'(y)$.

Once again Newton's Method is embodied in the formula

$$x_{n+1} = x_n - [f'(x_n)]^{-1} f(x_n) ,$$

where this time $f'(x_n)$ is regarded as an invertible linear operator.

5.1.3 Construction of the requisite approximate sets

Let E and F be two Banach spaces with $x_0 \in E$ and $U = \{x \in E : \| x - x_0 \| < q\}$, q some fixed number. If f is any mapping of U into F which is twice differentiable for each $x \in U$, we shall make the assumption about f

that there exists a constant k such that

$$\|f''(x)\| \leqslant k \quad \text{for all} \quad x \in U .$$

There are two consequences of this assumption which we shall find useful.

5.1.3.1 **Proposition** If $x \in U$ and $y \in U$ then

(1) $\|f'(y) - f'(x)\| \leqslant k\|y - x\|$;

(2) $\|f(y) - f(x) - f'(x)(y - x)\| \leqslant \frac{1}{2}k\|y - x\|^2$.

Background details and proofs are to be found in Dieudonné [2], Chapter VIII.

If A is a linear operator from E into F we define a measure of invertibility for A as follows:

$$d(A) = \inf \{\|Ax\| : \|x\| = 1\} .$$

This has the following properties:

5.1.3.2 **Proposition** If A^{-1} exists then

(1) $d(A) = \|A^{-1}\|^{-1}$;

(2) $d(A) - \|B\| \leqslant d(A + B)$ for any perturbation by a bounded linear operator B .

Proof (1) $d(A) = \inf\limits_{x \neq 0} \dfrac{\|Ax\|}{\|x\|}$

$$= \inf_{y \neq 0} \frac{\|y\|}{\|A^{-1}y\|}$$

$$= \frac{1}{\sup\limits_{y \neq 0} \dfrac{\|A^{-1}y\|}{\|y\|}} ,$$

i.e. $d(A) = \|A^{-1}\|^{-1}$.

(2) $$\|Ax\| \leqslant \|(A + B)x\| + \|Bx\|$$
$$\leqslant \|(A + B)x\| + \|B\| \quad \text{for} \quad \|x\| = 1$$

whence, taking infimums, $d(A) \leqslant \|(A + B)x\| + \|B\|$.
So, taking infimums again, $d(A) - \|B\| \leqslant d(A + B)$.

5.1.3.3 The Newton transformation

Suppose now that $f'(x)$ is invertible for each $x \in U$. Then we can define a mapping N of U into F as follows:

$$N(f,x) = x - (f'(x))^{-1} f(x) .$$

We shall call N the *Newton transformation* and the value $N(f,x) - x$ the *Newton increment* at the point x .

Provided the results stay within U it is possible to iterate the mapping. Let us write $N(n,f,x)$ for the value of the n-th iterate of $N(f,\cdot)$ at the point x . In other words, $N(n,f,x) = x_n$, the point x_n being determined recursively from the equation

$$x_{n+1} - x_n = - (f'(x_n))^{-1} f(x_n) , \quad x_0 = x .$$

5.1.3.4 The approximate sets

Now we are ready to define our system of approximate sets. For each sufficiently small t let

$$Z(t) = \{x \in E : \|(f'(x))^{-1} f(x)\| \leqslant t , d(f'(x)) \geqslant h(t)\} ,$$

where h is a positive function continuous at 0^+ to be determined later.

Then, clearly, $Z(0)$ furnishes the solution of $f(x) = 0$.

In order to be able to apply the Induction Theorem we require, for a suitable rate of convergence, the inclusion

$$Z(t) \subset B\Big(Z(\omega(t)),\ t\Big) .$$

Now suppose $x \in Z(t)$ and set $x' = x - (f'(x))^{-1}f(x)$ so that $\|x' - x\| \leqslant t$. Let us estimate $d(f'(x'))$. Clearly, by 5.1.3.2 (2) and 5.1.3.1 (1), we have that

$$d(f'(x')) \geqslant d(f'(x)) - \|f'(x') - f'(x)\| \geqslant h(t) - kt ,$$

i.e. $\|f'(x')\|^{-1} \leqslant \dfrac{1}{h(t) - kt}$, provided $h(t) - kt > 0$.

And since, by definition of x' ,

$$\|f(x')\| = \|f(x') - f(x) - f'(x)(x' - x)\| ,$$

5.1.3.1 (2) provides

$$\|f(x')\| \leqslant \tfrac{1}{2}kt^2 .$$

Hence $\|(f'(x'))^{-1}f(x')\| \leqslant \tfrac{1}{2}kt^2 \cdot \dfrac{1}{h(t) - kt}$.

Thus, in order to satisfy the desired inclusion, it will be sufficient to postulate the following inequalities:

$$h(t) - kt \geqslant h(\omega(t)) ;$$

$$\tfrac{1}{2}kt^2 \cdot \frac{1}{h(t) - kt} \leqslant \omega(t) .$$

5.1.4 Computation of the rate of convergence of Newton's Process

Our concern now becomes finding an interval T , a rate of convergence ω on T , and a positive function h , continuous at 0^+, on T such that

$$h(t) - kt \geqslant h(\omega(t))$$

and $\tfrac{1}{2}kt^2 \leqslant \omega(t)[h(t) - kt]$, for all $t \in T$.

We start by postulating equality in the first relationship:

$$h(t) - h(\omega(t)) = kt \ .$$

Upon iteration we obtain

$$\begin{aligned} h(t) - h(\omega(t)) &= kt \\ h(\omega(t)) - h(\omega^2(t)) &= k\omega t \\ h(\omega^2(t)) - h(\omega^3(t)) &= k\omega^2(t) \\ \vdots \qquad\quad \vdots \quad &\qquad \vdots \\ h(\omega^{n-1}(t)) - h(\omega^n(t)) &= k\omega^{n-1}(t) \ . \end{aligned}$$

From this it is clear that, if $b = \frac{1}{k} \lim\limits_{n\to\infty} h(\omega^n(t))$

$$= \frac{1}{k} h(0) \ ,$$

then $h(t) = k[b + \sigma(t)]$

or $h(\omega(t)) = k[b + \sigma(\omega(t))] = k[b + \sigma(t) - t]$,

where σ is the cumulative error function (as in 1.5).

Our task reduces thus to finding a rate of convergence ω such that

$$\frac{1}{2}t^2 \leqslant \omega(t)[b + \sigma(t) - t] \quad \text{for all} \quad t \in T \ .$$

In order to compute the rate of convergence we require some information about quadratic polynomials.

5.1.4.1 Lemma Let f be a quadratic of the form $f(x) = (x - p)^2 - d$ with $p > 0$, $d > 0$. If $r < p$ and if the Newton transformation of r results in a positive increment x , then the increment at the following step of Newton's Process is given by

$$\frac{x^2}{2(x^2 + d)^{1/2}} .$$

Proof The increment at r has the form

$$-\frac{f(r)}{f'(r)} = \frac{1}{2}\left(p - r - \frac{d}{p - r}\right) .$$

Suppose now that

$$\frac{1}{2}\left(p - r - \frac{d}{p - r}\right) = x > 0 .$$

We are to compute the increment $-\dfrac{f(r + x)}{f'(r + x)}$.

Writing u for $p - r$ it is easily shown that

$$-\frac{f(r + x)}{f'(r + x)} = \frac{1}{2}\left[\frac{1}{2}\left(u + \frac{d}{u}\right) - \frac{d}{\frac{1}{2}\left(u + \frac{d}{u}\right)}\right] .$$

And, of course, $\frac{1}{2}\left(u - \frac{d}{u}\right) = x$. Now, since

$$\left(u + \frac{d}{u}\right)^2 = \left(u - \frac{d}{u}\right)^2 + 4d ,$$

we have $\frac{1}{4}\left(u + \frac{d}{u}\right)^2 = x^2 + d$.

Using this we obtain

$$-\frac{f(r + x)}{f'(r + x)} = \frac{x^2}{2(x^2 + d)^{1/2}} ,$$

which gives the desired result.

We define

$$\omega(x) = \frac{x^2}{2(x^2 + d)^{1/2}}$$

and note that this function is a rate of convergence since $\sum_{n=1}^{\infty} \omega^n(x)$ converges by comparison with $\sum_{n=1}^{\infty} x^{2^n}$.

Now let us determine the cumulative error function corresponding to this ω . First we investigate the partial sums.

5.1.5 Formula for the partial sums of the cumulative error function in Newton's Process

The case $d = 0$ is straightforward since then the function is linear. Let us examine the case $d > 0$. We require two preliminary results concerning a recursively defined sequence.

5.1.5.1 Lemma If $x_0 > 1$ and a sequence x_n is defined by the relation

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right) ,$$

then

$$x_n = \frac{(x_0 + 1)^{2^n} + (x_0 - 1)^{2^n}}{(x_0 + 1)^{2^n} - (x_0 - 1)^{2^n}} .$$

Proof We look for solutions of the form $x_n = \dfrac{u_n}{v_n}$.

The relation to be satisfied becomes

$$\frac{u_{n+1}}{v_{n+1}} = \frac{1}{2}\left(\frac{u_n}{v_n} + \frac{v_n}{u_n}\right) .$$

This can be written as

$$\frac{u_{n+1}}{v_{n+1}} = \frac{1}{2}\,\frac{u_n{}^2 + v_n{}^2}{u_n v_n}$$

$$= \frac{(u_n + v_n)^2 + (u_n - v_n)^2}{(u_n + v_n)^2 - (u_n - v_n)^2} .$$

Now we set $u_n + v_n = p_n$, $u_n - v_n = q_n$, and reformulate the relation as follows:

$$\frac{p_{n+1} + q_{n+1}}{p_{n+1} - q_{n+1}} = \frac{p_n{}^2 + q_n{}^2}{p_n{}^2 - q_n{}^2} .$$

This will be satisfied if we set $p_{n+1} = p_n{}^2$ and $q_{n+1} = q_n{}^2$. Hence

$$p_{n+1} = p_n^2 = p_{n-1}^4 = p_{n-2}^8 = \ldots\ldots = p_0{}^{2^{n+1}}$$

and $$q_{n+1} = q_n^2 = q_{n-1}^4 = q_{n-2}^8 = \ldots\ldots = q_0{}^{2^{n+1}} ,$$

giving

$$x_{n+1} = \frac{p_0{}^{2^{n+1}} + q_0{}^{2^{n+1}}}{p_0{}^{2^{n+1}} - q_0{}^{2^{n+1}}}$$

or $$x_n = \frac{p_0{}^{2^n} + q_0{}^{2^n}}{p_0{}^{2^n} - q_0{}^{2^n}} ,$$

for a suitable choice of p_0 and q_0 . A possible choice is to take p_0 and q_0 such that

$$p_0 + q_0 = x_0 \quad \text{and} \quad p_0 - q_0 = 1 .$$

It is easily seen that this gives rise to $\frac{u_0}{v_0} = x_0$. With this choice of p_0 and q_0 we are led to the desired formula

$$x_n = \frac{(x_0 + 1)^{2^n} + (x_0 - 1)^{2^n}}{(x_0 + 1)^{2^n} - (x_0 - 1)^{2^n}} .$$

5.1.5.2 <u>Lemma</u> Suppose that $y_0 > d^{\frac{1}{2}}$ and that the sequence y_n is defined by the recursive formula

$$y_{n+1} = \frac{1}{2}\left(y_n + \frac{d}{y_n}\right) ,$$

then $$y_n = d^{1/2}\frac{(y_0 + d^{1/2})^{2^n} + (y_0 - d^{1/2})^{2^n}}{(y_0 + d^{1/2})^{2^n} - (y_0 - d^{1/2})^{2^n}} .$$

Proof If we set $y_n = d^{1/2}x_n$ then x_n satisfies the relation

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right)$$

and the required result follows from 5.1.5.1 in an elementary way.

Now we are ready to derive the formula for the partial sums of the cumulative error function under discussion.

5.1.5.3 **Theorem** Let $d > 0$. If $\omega(x) = \frac{x^2}{2(x^2 + d)^{1/2}}$, then for each $n \in \mathbb{N}$ and each $x > 0$ we have

$$\sum_{r=0}^{n-1} \omega^r(x) = x + (x^2 + d)^{1/2} - d^{1/2}\frac{[d^{1/2} + (x^2 + d)^{1/2}]^{2^n} + x^{2^n}}{[d^{1/2} + (x^2 + d)^{1/2}]^{2^n} - x^{2^n}} .$$

Proof Let f be the function defined for real x by the formula

$$f(x) = x^2 - d \quad \text{(of the type in 5.1.4.1)} .$$

Consider a point $x_0 > d^{1/2}$ and the Newton Process for f starting at x_0. Since

$$\frac{f(x)}{f'(x)} = \frac{1}{2}\left(x - \frac{d}{x}\right),$$

the Newton Process transforms a point $z \neq 0$ into the point

$$N(z) = z - \frac{f(z)}{f'(z)} = \frac{1}{2}\left(z + \frac{d}{z}\right) .$$

Now suppose that x_0 is such that

$$x_0 - N(x_0) = x \ ,$$

then, by 5.1.4.1, $N(x_0) - N^2(x_0) = \omega(x)$;

$$\begin{aligned} N^2(x_0) - N^3(x_0) &= \omega^2(x) ; \\ \vdots \qquad\quad \vdots \qquad &\qquad \vdots \\ N^{n-1}(x_0) - N^n(x_0) &= \omega^{n-1}(x) \ . \end{aligned}$$

Adding, we obtain $x_0 - N^n(x_0) = x + \omega(x) + \ldots + \omega^{n-1}(x)$.

Thus, applying Lemma 5.1.5.2 we now have

$$N^n(x_0) = d^{1/2} \frac{(x_0 + d^{1/2})^{2^n} + (x_0 - d^{1/2})^{2^n}}{(x_0 + d^{1/2})^{2n} - (x_0 - d^{1/2})^{2n}}$$

$$= d^{1/2} \frac{\left[\dfrac{d^{1/2} + (x^2 + d)^{1/2}}{x}\right]^{2^n} + 1}{\left[\dfrac{d^{1/2} + (x^2 + d)^{1/2}}{x}\right]^{2n} - 1} ,$$

which proves the theorem.

5.1.6 The cumulative error function in Newton's Process

As a direct consequence of the result proved in 5.1.5.3 above we obtain

$$\sigma(x) = \lim_{n\to\infty} \sum_{r=0}^{n} \omega^r(x) = x + (x^2 + d)^{1/2} - d^{1/2} \ .$$

5.1.7 Convergence of the Newton Process

At this stage of the discussion we have all the necessary tools at our disposal for verifying the convergence of the Newton Process described above.

Let E, F, x_0, U, q, f, k and d be as in 5.1.3. Suppose we are given the initial estimates

$$d(f'(x_0)) = d_0 > 0$$

and $$\|(f'(x_0))^{-1} f(x_0)\| = t_0 ,$$

where $\frac{d_0}{k} - 2t_0 \geqslant 0$.

5.1.7.1 **Theorem** If $d = \frac{d_0}{k}\left(\frac{d_0}{k} - 2t_0\right)$ and $q \geqslant \frac{d_0}{k} - d^{1/2}$, then the Newton Process starting at x_0 , with rate of convergence $\omega(t) = \frac{t^2}{2(t^2 + d)^{1/2}}$, converges to a point x such that $f(x) = 0$.

The following estimates are yielded:

$$\|x - x_0\| \leqslant \frac{d_0}{k} - d^{1/2} ;$$

$$\|x_{n+1} - x_n\| \leqslant \omega^n(t_0) .$$

Proof In view of the discussion in 5.1.4 it suffices

(i) to prove that

$$\frac{1}{2}t^2 \leqslant \omega(t)[b + \sigma(t) - t] ,$$

where $$b = \frac{1}{k}h(0) ,$$

then (ii) to set

$$h(t) = k[b + \sigma(t)]$$

and verify the inequality

$$d_0 \geqslant h(t_0) .$$

For (i) set $b = d^{1/2}$, then by definition of $\omega(t)$,

$$\frac{1}{2}t^2 = \omega(t)(t^2 + d)^{1/2} ,$$

and it follows from 5.1.6 that

$$\frac{1}{2}t^2 = \omega(t)[d^{1/2} + \sigma(t) - t] \quad \text{for all} \quad t > 0 .$$

For (ii) note that

$$\begin{aligned} h(t_0) &= k[d^{1/2} + \sigma(t_0)] \\ &= k[t_0 + (t_0^2 + d)^{1/2}] . \end{aligned}$$

This gives

$$\begin{aligned} h(t_0) &= k\left[t_0 + \left(t_0^2 + \left(\frac{d_0}{k}\right)^2 - 2\frac{d_0}{k}t_0\right)^{1/2}\right] \\ &= kt_0 + (k^2t_0^2 + d_0^2 - 2d_0kt_0)^{1/2} . \end{aligned}$$

Now application of the Induction Theorem (2.3.2) yields the existence of an $x \in Z(0)$ (thus a solution of $f(x) = 0$) within distance $\sigma(t_0)$ from x_0.

The Newton Process is meaningful since

$$\begin{aligned} \sigma(t_0) &= t_0 + (t_0^2 + d)^{1/2} - d^{1/2} \\ &= t_0 + \left(\frac{d_0}{k} - t_0\right) - d^{1/2} \\ &= \frac{d_0}{k} - d^{1/2} , \end{aligned}$$

i.e. $\|x - x_0\| \leqslant q$.

For $f(x) = \frac{1}{2}kx^2 + d_0x + d_0t_0$ and $x_0 = 0$ the estimates are attained, *viz.*

$$d(f'(x_0)) = d_0 \quad \text{and} \quad \|(f'(x_0))^{-1} f(x_0)\| = t_0 .$$

5.2 Moser's Theorem – a modification of Newton's Method

Pták has used his Induction Theorem to simplify the proof of Moser's Theorem concerning a modification of

Newton's Method [16]. Unfortunately, this work has very little aesthetic appeal, the algebraic detail involved being cumbersome to say the least. For this reason it is probably best overlooked. Suffice it to say that Moser's Theorem provides an opportunity for an application of the Induction Theorem in its form 2.3.4.

Section 6

ASSORTED APPLICATIONS OF THE INDUCTION THEOREM

The main advantage of the non-discrete approach seems to be that, by separating the analysis involved in a problem from the construction of the necessary iterative process, the Induction Theorem not only yields considerable simplification of proofs but also reveals more clearly the essential features of the problem. This fact is reasonably well illustrated by the applications of the theorem dealt with in the preceding section. In this section our aim will be to provide further illustrations by applying the Induction Theorem to a few quite diverse problems in analysis: finding eigenvalues of almost decomposable matrices; proving a Selection Theorem; factorization in Banach algebras without identity; strict irreducibility of $*$-representations of B$*$-algebras.

6.1 Eigenvalues of almost decomposable matrices

In 1964 Fiedler and Pták in a joint paper [5] discussed the problem of almost decomposable operators. They treated several iteration processes for eigenvalues of almost decomposable matrices. The methods they used were of a classical nature. In 1976, however, Pták re-examined the problem and found the non-discrete induction approach to be superior. We shall discuss aspects of his paper [12].

6.1.2 Formulation of the problem

Let F be a normed space, C a bounded linear operator on F, and let a be a complex number. In addition, let u be an element of F and v a bounded linear functional on F. Then define E as the direct product of a one-dimensional space and F. The elements of E will thus be pairs $x = (x_1, x_2)$ where x_1 is a complex number and $x_2 \in F$. Now define a linear operator A on E by the matrix

$$A = \begin{pmatrix} a & v \\ u & C \end{pmatrix} .$$

This is to be understood as follows: the equation $y = Ax$ is equivalent to

$$y_1 = x_1 a + \langle x_2, v \rangle ,$$

$$y_2 = x_1 u + C x_2 \quad .$$

Solving this sytem for x_1 to obtain $x_1 = \dfrac{y_1 - \langle C^{-1} y_2, v \rangle}{a - \langle C^{-1} u, v \rangle}$, we arrive at the following result: suppose that C^{-1} exists, then A is invertible if and only if

$$a - \langle C^{-1} u, v \rangle \neq 0 .$$

Suppose now that the vector u and the functional v are small, making A "almost diagonal". It is to be expected then that A will have an eigenvalue close to a. To verify this we must find λ satisfying the characteristic equation

$$a - \lambda = \langle (C - \lambda)^{-1} u, v \rangle .$$

If we set $z = \lambda - a$ and $D = C - a$, the problem reduces to solving for z the equation

$$-z = \langle (D - z)^{-1}u,v \rangle .$$

6.1.3 Solution of the problem using a linear rate of convergence

In order to define a system of approximate sets we introduce the function

$$g(z) = \langle (D - z)^{-1}u,v \rangle ,$$

defined in a neighbourhood of the origin. Conditions are to be found which will ensure the existence of a solution of the equation

$$z + g(z) = 0 ,$$

i.e. of
$$-z = \langle (D - z)^{-1}u,v \rangle .$$

As in 5.1.3 we shall use the measure of invertibility

$$d(T) = \inf \{\|Tx\| : \|x\| = 1\} ,$$

so that $d(T) = \|T^{-1}\|^{-1}$ if T is invertible. Note also that $d(T - \lambda) \geqslant d(T) - \lambda$ for any complex λ.

For $t > 0$ set

$$Z(t) = \{z : |z + g(z)| \leqslant t , d(D - z) \geqslant h(t)\} ,$$

where h is a positive function to be chosen later.

Suppose now that $z \in Z(t)$. Then

$$d(D + g(z)) \geqslant d(D - z) - |z + g(z)| \geqslant h(t) - t .$$

In addition, for $z' = -g(z)$, assuming both $(D - z)^{-1}$ and $(D - z')^{-1}$ exist, by a standard result of operator theory

(Taylor [19], Chapter 5), we have

$$z' + g(z') = (z' - z)\langle (D - z)^{-1}(D - z')^{-1}u,v\rangle .$$

Hence

$$|z' + g(z')| \leqslant t[h(t)]^{-1}[h(t) - t]^{-1}\|u\|\|v\| .$$

We require, for a suitable rate of convergence ω, the inclusion

$$Z(t) \subset B\Big(Z(\omega(t)),t\Big) .$$

So, since $z \in Z(t)$, we must show that $z' \in Z(\omega(t))$. For this the following inequalities are sufficient:

(1) $t[h(t)]^{-1}[h(t) - 1]^{-1}\|u\|\|v\| \leqslant \omega(t)$;

(2) $h(t) - t \geqslant h\big(\omega(t)\big)$.

Our problem is thus reduced to finding a small function ω and a positive function h, each defined for small positive t and satisfying (1) and (2) above. In order to be able to apply the Induction Theorem we shall also require the existence of $t_0 > 0$ such that $Z(t_0) \neq \emptyset$. It will be shown that this is realized by the inclusion $0 \in Z(t_0)$ or

$$|g(0)| \leqslant t_0 \quad \text{and} \quad \|D^{-1}\|^{-1} \geqslant h(t_0) .$$

Now, if ω is a small function with cumulative error function σ, then for any b, the function $h(t) = b + \sigma(t)$ satisfies the functional equation

$$h(t) - t = h\big(\omega(t)\big) ,$$

i.e. (2) above. Hence, using a telescoping process as in 5.1.4, it is conceivable that, for a suitable choice of b

and ω , the function $h(t) = b + \sigma(t)$ will also satisfy requirement (1). For this choice of h the inclusion $0 \in Z(t_0)$ is equivalent to

$$|g(0)| \leqslant t_0 \quad \text{and} \quad \|D^{-1}\|^{-1} \geqslant b + \sigma(t_0) .$$

For convenience let us replace $\|D^{-1}\|^{-1}$ by γ . If σ is non-decreasing we may write the above two conditions in the form of a single inequality

$$(3) \quad \sigma(|g(0)|) \leqslant \gamma - b .$$

For the sake of simplicity we try to satisfy our system of functional inequalities initially by means of a linear rate of convergence. Take ω in the form $\omega(t) = \alpha t$ where α , $0 < \alpha < 1$, is to be chosen later. It follows that

$$h(t) = b + \frac{1}{1 - \alpha} \cdot t .$$

Hence the conditions to be satisfied ((1) and (3) above) are as follows:

$$(1') \quad \|u\|\|v\| \leqslant \alpha\left(b + \frac{1}{1 - \alpha} \cdot t\right)\left(b + \frac{\alpha}{1 - \alpha} \cdot t\right);$$

$$(3') \quad \frac{1}{1 - \alpha} \cdot |g(0)| \leqslant \gamma - b .$$

Here, for convenience, we shall replace $\|u\|\|v\|$ by β . Then, to satisfy (1'), $\beta \leqslant \alpha b^2$ is sufficient. Obviously the easiest possibility to test is $\beta = \alpha b^2$.

So we set

$$b = \sqrt{\frac{\beta}{\alpha}} ,$$

leaving α as the parameter to be chosen.

Now (3') cannot be satisfied unless

$$\gamma - \sqrt{\frac{\beta}{\alpha}} = \gamma - b \geqslant 0 \ .$$

It is thus necessary to assume $\gamma^2 > \beta$. Since

$$|g(0)| = |\langle D^{-1}u,v\rangle| \leqslant \|D\|^{-1}\|u\|\|v\| = \frac{\beta}{\gamma} \ ,$$

(3') will be satisfied if $\frac{\beta}{\gamma} \leqslant (1-\alpha)\left(\gamma - \sqrt{\frac{\beta}{\alpha}}\right)$,

i.e. if $1 \leqslant (1-\alpha)\left[\frac{\gamma^2}{\beta} - \sqrt{\frac{\gamma^2}{\beta}} \cdot \frac{1}{\sqrt{\alpha}}\right]$.

Now write $\sqrt{\frac{\gamma^2}{\beta}} \cdot \frac{1}{\sqrt{\alpha}} = \frac{\gamma^2}{\beta} \cdot \frac{1}{\delta}$,

replacing the parameter α by δ in such a way that

$$\alpha = \frac{\beta}{\gamma^2} \cdot \delta^2 \ , \text{ where we shall require } \delta > 1 \ .$$

Then (3') will be satisfied if

$$\frac{\gamma^2}{\beta} \geqslant \delta^2 + \frac{\delta}{\delta - 1} \ .$$

Denote by δ_0 the (unique) point where $\delta^2 + \frac{\delta}{\delta - 1}$, $\delta > 1$, assumes its minimum value m_0 . We show that it suffices to postulate $\frac{\gamma^2}{\beta} \geqslant m_0$ and set $\alpha = \frac{\beta}{\gamma^2} \cdot \delta_0$.

If $\frac{\gamma^2}{\beta} \geqslant m_0$, then

$$1 \geqslant \frac{\beta}{\gamma^2} \cdot \delta_0^{\ 2} + \frac{\beta}{\gamma^2} \cdot \frac{\delta_0}{\delta_0 - 1} \ ,$$

so that $\alpha = \frac{\beta}{\gamma^2} \cdot \delta_0^{\ 2} < 1$.

For this α the rate of convergence $\omega(t) = \alpha t$ satisfies all our requirements.

The Induction Theorem (2,3,2) ensures the existence of $z \in Z(0)$. Thus

$$|z + g(z)| = 0 \quad \text{and} \quad d(D - z) \geqslant h(0) = b \ .$$

So $|z| = |g(z)| = |\langle (D - z)^{-1}u, v\rangle|$

$$\leqslant \|(D - z)^{-1}\| . \|u\| . \|v\| \ .$$

Now $d(D - z) = \dfrac{1}{\|(D - z)^{-1}\|}$; hence

$$|z| \leqslant \frac{1}{b} \cdot \beta \, .$$

But we set $b = \sqrt{\dfrac{\beta}{\alpha}}$, so

$$|z| \leqslant \sqrt{\beta} \, . \, \sqrt{\alpha},$$

i.e. $|\lambda - a| \leqslant \sqrt{\|u\| . \|v\|} \, . \, \sqrt{\alpha} < \sqrt{\|u\| . \|v\|} \qquad (|\alpha| < 1)$.

So, if u and v are small, there is an eigenvalue λ near a (as predicted in 6.1.2).

Pták gives the inequality

$$|\lambda - a| \leqslant \frac{1}{1 - \left(\frac{\beta}{\gamma^2}\right)\delta_0^{\,2}} |\langle D^{-1} u, v \rangle| \ .$$

Minimizing $\delta^2 + \dfrac{\delta}{\delta - 1}$ reduces to solving the cubic $2\delta^3 - 4\delta^2 + 2\delta - 1 = 0$. Approximate values of m_0 and δ_0 are

$$\delta_0 = 1{,}56 \ , \qquad m_0 = 5{,}22 \ ,$$

and this seems to be the best we can do using a linear rate of convergence.

6.1.4 Solution of the problem using a more refined rate of convergence

We show now that by imposing further restrictions on the family $Z(\cdot)$ we can obtain a more precise result.

Consider firstly $|g(z)|$:

$$g(z) = \langle (D - z)^{-1} u, v \rangle$$

whence $|g(z)| \leqslant \frac{\|u\|\|v\|}{d(D - z)} \leqslant \frac{\beta}{h(t)}$ if $z \in Z(t)$.

Now $d(D - z') \geqslant d(D) - |z'| = d(D) - |g(z)|$,

so we have

$$d(D - z') \geqslant \gamma - \frac{\beta}{h(t)} .$$

Remembering that $d(D - z') \geqslant h(t) - t$ and that previously we required $h(t) - t \geqslant h(\omega(t))$, it is natural to impose the following additional requirement:

$$(4) \qquad \gamma - \frac{\beta}{h(t)} \geqslant h(\omega(t)) .$$

Now we can define a more precise approximate set as follows:

$$Z(t) = \{z : |z + g(z)| \leqslant t , \quad d(D - z) \geqslant h(t), \quad \gamma - |z| \geqslant h(t)\} .$$

If conditions (1) and (2) of 6.1.3 and (4) above are satisfied , the inclusion

$$Z(t) \subset B\Big(Z(\omega(t)), t\Big)$$

will hold.

As in 6.1.3 we shall satisfy (2) by postulating equality:

$$(2') \qquad h(t) - t = h(\omega(t)) .$$

To satisfy (4) it will be sufficient to have

$$h(t) - t - \gamma + \frac{\beta}{h(t)} = 0$$

$$\text{or} \qquad [h(t)]^2 - (t + \gamma)h(t) + \beta = 0 .$$

Solving for $h(t)$ we obtain

$$(4') \quad h(t) = \frac{\gamma + t}{2} + \sqrt{\left(\frac{\gamma + t}{2}\right)^2 - \beta} \ ,$$

so, in order to ensure the existence of a solution for small positive t , it will be necessary to impose the restriction $\gamma^2 > 4\beta$.

Now, to obtain the rate of convergence corresponding to the function

$$\sigma(t) = h(t) - b \ ,$$

all we need do is substitute the expression for $h(t)$ given by (4') into (2'). A little elementary algebraic manipulation yields

$$\omega(t) = t \cdot \frac{\gamma + t - v}{\gamma - t + v} \ .$$

This, we observe, is the rate of convergence mentioned in 1.2.3.

Now all that remains is to verify (1). This is straightforward since

$$\beta = h(t) \, . \, [\gamma + t - h(t)] \ ;$$

hence $t[h(t)]^{-1}[h(t) - t]^{-1}\beta = t \, . \, \dfrac{\gamma + t - h(t)}{h(t) - t} = \omega(t)$.

Now, using the definition of $Z(\cdot)$, it is easy to prove that $0 \in Z\left(\frac{\beta}{\gamma}\right)$. So, it follows from the Induction Theorem that $Z(0) \neq \emptyset$. But clearly, $z \in Z(0)$ implies $z + g(z) = 0$, so that $z + a$ is an eigenvalue of A . Also,

$$\gamma - |z| \geqslant h(0) = \tfrac{1}{2}\left(\gamma + \sqrt{\gamma^2 - 4\beta}\right) ,$$

so that
$$|z| \leqslant \tfrac{1}{2}\left(\gamma - \sqrt{\gamma^2 - 4\beta}\right) ,$$

i.e.
$$|\lambda - a| \leqslant \tfrac{1}{2}\left(\gamma - \sqrt{\gamma^2 - 4\beta}\right) .$$

6.2 A Selection Theorem

6.2.1 Definition Given A and F topological spaces and for each $a \in A$ there exists a non-void set $H(a) \subset F$. A *selection* from (associated with) $H(\cdot)$ is a mapping $f : A \to F$ such that $f(a) \in H(a)$ for each $a \in A$.

We shall require $H(\cdot) \in \mathscr{C}$, a class of sets closed under intersections with spherical neighbourhoods.

6.2.2 Definition The mapping $a \longmapsto H(a)$ is called *continuous* if for each $a_0 \in A$, given V an open set such that $V \cap H(a_0) \neq \emptyset$, there exists an open neighbourhood W of a_0 such that, for each $a \in W$, $V \cap H(a) \neq \emptyset$.

Note that this idea of continuity coincides with ordinary continuity when H is a function, i.e. $H(a)$ is a point for all $a \in A$.

Denote by $\mathscr{P}(F)$ the power set of F.

6.2.3 Definition Given $H(\cdot) \in \mathscr{C}$, let G be a set-valued mapping, $G : A \to \mathscr{P}(F)$, such that $H(a) \cap G(a) \neq \emptyset$ for all $a \in A$. The mapping G will be said to be *pseudo-continuous* if $a \longmapsto G(a)$ satisfies the following condition: for each $a_0 \in A$ and each $g_0 \in G(a_0)$, there exists an open set $Q(a_0) \ni g_0$ and a neighbourhood U of a_0 such that, for each $a \in U$, $Q(a_0) \subset G(a)$.

6.2.4 Lemma If A and F are topological spaces and

$G : A \to \mathcal{P}(F)$ is pseudo-continuous, then continuity of $a \mapsto H(a)$ implies continuity of $a \mapsto H(a) \cap G(a)$.

Proof Let V be an open subset of F and let $a_0 \in A$ be such that $V \cap H(a_0) \cap G(a_0) \neq \emptyset$. Take a point $g_0 \in V \cap H(a_0) \cap G(a_0)$. Then by pseudo-continuity of G there exists a neighbourhood U of a_0 and an open set $Q(a_0) \ni g$ such that $Q(a_0) \subset G(a)$ for $a \in U$. So $g_0 \in H(a_0) \cap V \cap Q(a_0)$ and since $V \cap Q(a_0)$ is open it follows from the continuity of $a \mapsto H(a)$ that there exists a neighbourhood W of a_0 such that $H(a) \cap V \cap Q(a_0) \neq \emptyset$ for all $a \in W$.

Suppose now that $a \in U \cap W$. Since $a \in W$, $H(a) \cap Q(a_0) \cap V \neq \emptyset$. Furthermore, $a \in U$ implies $H(a) \cap Q(a_0) \cap V \subset H(a) \cap G(a) \cap V$. This last intersection is thus non-void for all a in the neighbourhood $U \cap W$ of a_0 , showing that the mapping $a \mapsto H(a) \cap G(a)$ is continuous.

6.2.5 The Approximate Selection Property

The triple $(A,F,\mathscr{C})$ will be said to possess the *approximate selection property* if, given any continuous relation H from A into F such that $H(a) \neq \emptyset$ and $H(a) \in \mathscr{C}$ for all $a \in A$, and given any $t > 0$, there exists a continuous selection f from the sets $B(H(a),t)$.

6.2.6 Requisite distance functions

Given A a topological space and F a metric space, we denote by $C(A,F)$ the space of all continuous functions from A to F.

6.2.6.1 **Definition** If x and y are elements of $C(A,F)$, the *distance between* x and y is

$$d(x,y) = \sup \{d(x(a), y(a)) : a \in A\} .$$

6.2.6.2 **Definition** If $z \in C(A,F)$, the distance between $z(a)$ and $H(a)$ is

$$d(z(a), H(a)) = \inf \{t : z(a) \in B(H(a),t)\}.$$

This is, of course, the usual notion of distance from a point to a set.

We shall also employ the notion of semi-continuity as defined in 2.2.1.

Having developed the necessary background material we are now ready to see how the Induction Theorem may be used to prove a Selection Theorem (Pták [10]).

6.2.7 **Theorem** Let A be a topological space and F a complete metric space such that $(A,F,\mathscr{C})$ possesses the approximate selection property. For each $a \in A$ let $H(a) \neq \emptyset$ be a closed subset of F, $H(a) \in \mathscr{C}$, then continuity of $H(\cdot)$ implies the existence of a continuous selection.

Proof Let $z \in C(A,F)$ be such that

$$d(z(a), H(a)) < 1 \quad \text{for each} \quad a \in A .$$

Such a function exists by the approximate selection property.

Denote by E the subspace of $C(A,F)$ whose elements x are such that $d(x,z) < \infty$. Then, since F is a complete metric space, E is a complete metric space.

Define, for each sufficiently small t, a set $Z(t) \subset E$ by

$$Z(t) = \{x \in E : d(x(a), H(a)) < t \text{ for all } a \in A\} .$$

Note that $Z(\cdot)$ forms a contracting family and that, if $Z(0) \neq \emptyset$, then any element of $Z(0)$ is a selection.

We shall employ the Induction Theorem in its semicontinuity form (2.3.1). Firstly we show that

$$(*) \qquad B\left(x, \tfrac{3}{2}t\right) \cap Z\left(\tfrac{1}{2}t\right) \neq \emptyset \text{ for all } x \in Z(t),$$

in order to prove semicontinuity of $Z(\cdot)$ with $q(t) = \frac{2}{3}t$, $p(t) = \frac{1}{3}t$ and $\omega(t) = \frac{1}{2}t$.

Let $x \in Z(t)$ and, for each a, set

$$G(a) = \{g \in F : d(g,x(a)) < t\} ,$$

$$M(a) = H(a) \cap G(a) \in \mathscr{C} .$$

Then, since $x \in Z(t)$, $M(a) \neq \emptyset$. And, G is pseudocontinuous, so according to our lemma (6.2.4), $a \mapsto M(a)$ is continuous. It follows from the approximate selection property that there exists $x' \in C(A,F)$ such that

$$x'(a) \in B\left(M(a), \tfrac{1}{2}t\right) \text{ for each } a \in A ,$$

i.e. $d(x'(a), M(a)) < \frac{1}{2}t$ for each $a \in A$.

Hence for each $a \in A$, there exists $h(a) \in M(a) \subset G(a)$

such that $d(x'(a), h(a)) < \frac{1}{2}t$. It follows that

$$\begin{aligned} d(x'(a), x(a)) &\leq d(x'(a), h(a)) + d(h(a), x(a)) \\ &< \tfrac{1}{2}t + t = \tfrac{3}{2}t , \end{aligned}$$

which shows that $x' \in B\left(x, \frac{3}{2}t\right)$ and $x' \in E$. In addition,

$$d(x'(a), H(a)) < \tfrac{1}{2}t \quad \text{for all} \quad a \in A .$$

Thus we have $x' \in B\left(H(a), \frac{1}{2}t\right)$, i.e. $x' \in Z\left(\frac{1}{2}t\right)$ and (*) is established.

Since $Z\left(\frac{1}{2}t\right) \neq \emptyset$ the Induction Theorem (2.3.1) tells us that $Z(0) \neq \emptyset$ and the proof is complete.

6.3 A Factorization Theorem

6.3.1 Definition Let A be an algebra over the real or complex field and M a linear space over the same field. M is said to be a *left* A-*module* if a mapping $(a,m) \to am$ of $A \times M$ into M is specified which satisfies the following axioms:

for each fixed $a \in A$, the mapping $m \to am$ is linear on M ;

for each fixed $m \in M$, the mapping $a \to am$ is linear on A ;

$$a_1(a_2m) = (a_1a_2)m \qquad (a_1,a_2 \in A , \quad m \in M) .$$

Let A be a Banach algebra without a unit element. We shall denote by A' the Banach algebra obtained by adjunction of an identity to A . Its elements are the formal sums $x + \lambda$ where $x \in A$ and λ is a complex number, the norm being defined as $\|x + \lambda\| = \|x\| + |\lambda|$. Let F be a Banach space which is a left A-module with the property that $\|ax\| \leqslant \|a\|\|x\|$ for all $a \in A$ and $x \in F$.

6.3.2 **Definition** (A,F) possesses a *weak approximate unit* of norm β if, for each $a \in A$, each $y \in F$ and each $\varepsilon > 0$, there exists an element $e \in A$ such that

$$\|e\| \leqslant \beta \quad (\beta \text{ minimal}) , \quad \|ea - a\| < \varepsilon , \quad \|ey - y\| < \varepsilon .$$

6.3.3 **Lemma** Let A be a Banach algebra without a unit element and let A' be the Banach algebra obtained by adjunction of an identity to A . Let β be a positive number. If $x \in A$ and $\|x\| \leqslant \beta$, set $c(x) = \frac{\beta + 1}{\beta} - \frac{1}{\beta}x \in A$. Then

(i) $c(x) - 1 = \frac{1}{\beta}(1 - x)$,

(ii) $[c(x)]^{-1}$ exists in A' and $\|[c(x)]^{-1}\| \leqslant \beta$,

(iii) $[c(x)]^{-1} - \frac{\beta}{\beta + 1} \in A$ and

$$\left\|[c(x)]^{-1} - \frac{\beta}{\beta + 1}\right\| \leqslant \frac{\beta^2}{\beta + 1} .$$

Proof (i) $c(x) = \frac{\beta + 1}{\beta}(1 - z)$ where $z = \frac{1}{\beta + 1}x$, so that $\|z\| \leqslant \frac{\beta}{\beta + 1} < 1$.

It follows that

(ii) $[c(x)]^{-1} = \frac{\beta}{\beta + 1} (1 + z + z^2 + \ldots)$

whence $\|[c(x)]^{-1}\| \leqslant \beta$,

and (iii) $\left\|[c(x)]^{-1} - \frac{\beta}{\beta + 1}\right\| \leqslant \frac{\beta^2}{\beta + 1}$.

Now let us proceed to the theorem (Pták [10]).

6.3.4 **Theorem** Let A be a Banach algebra without a unit element and F a Banach space which is a left A-module. If (A,F) possesses a weak approximate unit, e, of norm β, then for each $y \in F$ and each sufficiently small $\varepsilon > 0$ there exists an $a \in A$ with $\|a\| \leqslant \beta^2 + 2\beta$, $z \in \overline{A'y}$ (closure in F) and $\|z - y\| < \varepsilon$, such that $y = az$.

Proof Let us first show that it suffices to prove the following slightly weaker result: for all sufficiently small $\alpha > 0$ there exists a factorization $y = az$ with $\|a\| \leqslant \beta^2 + 2\beta + \alpha$, $z \in \overline{A'y}$ and $\|z - y\| < \alpha$. To see this let us suppose that this result is proved and that $\varepsilon > 0$ is given. Let $y = a_0 z_0$ be a factorization of the type just mentioned, subject to the condition $\alpha < \beta^2 + 2\beta$. Set $\delta = \dfrac{\beta^2 + 2\beta}{\beta^2 + 2\beta + \alpha}$ and $a = \delta a_0$, $z = \delta^{-1} z_0$.

Then $\|a\| \leqslant \delta\|a_0\|$, i.e. $\|a\| \leqslant \beta^2 + 2\beta$; furthermore

$$\|z - y\| = \|\delta^{-1} z_0 - y\| = \|(\delta^{-1} - 1)y + \delta^{-1}(z_0 - y)\| ,$$

i.e. $\|z - y\| \leqslant (\delta^{-1} - 1)\|y\| + \delta^{-1}\|z_0 - y\|$

$$< (\delta^{-1} - 1)\|y\| + 2\alpha ,$$

i.e. $\|z - y\| < \left(\dfrac{1}{\beta^2 + 2\beta}\|y\| + 2\right)\alpha$.

Clearly α can be chosen sufficiently small to have this last term $< \varepsilon$. So all we need do is prove the weaker result described above.

Fix $\varepsilon > 0$, $\varepsilon < \min\left\{1, \frac{\beta^2}{1+\beta}\right\}$.

Let E be the cartesian product

$$E = \{a \in A : \|a\| \leqslant \beta^2 + 2\beta + \varepsilon\} \times \overline{A'y} ,$$

with the norm of $x = (a,z)$ taken as

$$\|x\| = \max\left\{\|a\|, \frac{(\beta+1)^2}{\varepsilon}\|z\|\right\} .$$

Then E is a complete metric space, being a product of closed (therefore complete) metric spaces.

For each $t < \beta^2 + 2\beta + \varepsilon$ define $Z(t) \subset E$ in the following way:

a pair $x = (a,z) \in E$ belongs to $Z(t)$ if and only if

(i) $\|a\| \leqslant \beta^2 + 2\beta + \varepsilon - (\beta+1)t$,

(ii) there exists a number γ , $|\gamma| < \frac{(\beta+1)t}{\beta^2 + 2\beta + 2}$, such that $(a+\gamma)^{-1}$ exists (in A') ,

and (iii) $z = (a+\gamma)^{-1}y$.

Suppose now that $x \in Z(t)$ and let us prove that there exists an x' such that

$$(*) \quad x' \in B(x,t) \cap Z\left(\frac{\beta}{\beta+1}t\right) ,$$

in order to be able to apply the Induction Theorem (2.3.2) where $\omega(t) = \frac{\beta}{\beta+1}t$.

We shall show that it suffices to take $x' = (a',z')$ where

$$a' = [c(e)]^{-1}(a+\gamma) - \frac{\beta}{\beta+1}\gamma ,$$

$$z' = (a + \gamma)^{-1} c(e) y ,$$

and c is as in Lemma 6.3.3. The weak approximate unit e is chosen so as to have

$$\text{(A)} \quad \|(e - 1)a\| + \frac{\beta^2 + 2\beta + 2}{\beta + 1} |\gamma| < t ,$$

and $\text{(B)} \quad \|(a + \gamma)^{-1}\| \|(e - 1)y\| < \beta \frac{\varepsilon}{(\beta + 1)^2} t .$

Relation (A) is made possible by our choice of γ such that

$$|\gamma| < \frac{\beta + 1}{\beta^2 + 2\beta + 2} \cdot t \quad \text{((ii) above)} .$$

Now to verify the inclusion (*) it will be sufficient to prove the following facts:

(iv) $a' \in A$;

(v) $\|x' - x\| < t$;

(vi) $\|a'\| \leqslant \beta^2 + 2\beta + \varepsilon - \beta t$;

(vii) $\left(a' + \frac{\beta}{\beta + 1} \cdot \gamma\right)^{-1}$ exists in A' ;

(viii) $\left(a' + \frac{\beta}{\beta + 1} \cdot \gamma\right) z' = y$.

To prove (iv) it suffices to write a' in the form:

$$a' = [c(e)]^{-1}(a + \gamma) - \frac{\beta}{\beta + 1}(a + \gamma) + \frac{\beta}{\beta + 1} \cdot a$$

$$= \left([c(e)]^{-1} - \frac{\beta}{\beta + 1}\right)(a + \gamma) + \frac{\beta}{\beta + 1} \cdot a ,$$

and note that by the preceding lemma (6.3.3),

$$[c(e)]^{-1} - \frac{\beta}{\beta + 1} \in A .$$

To see that (v) is satisfied, write

$$\begin{aligned} a' - a &= ([c(e)]^{-1} - 1)(a + \gamma) + \left(1 - \frac{\beta}{\beta + 1}\right)\gamma \\ &= [c(e)]^{-1}(1 - c(e))(a + \gamma) + \left(1 - \frac{\beta}{\beta + 1}\right)\gamma \\ &= [c(e)]^{-1}\left[1 - \left(\frac{\beta + 1}{\beta} - \frac{e}{\beta}\right)\right](a + \gamma) + \left(1 - \frac{\beta}{\beta + 1}\right)\gamma \\ &= [c(e)]^{-1} \cdot \frac{1}{\beta}(e - 1)(a + \gamma) + \frac{1}{\beta + 1}\gamma . \end{aligned}$$

Hence, since $\|[c(e)]^{-1}\| \leqslant \beta$ by (ii) of 6.3.3,

$$\begin{aligned} \|a' - a\| &\leqslant \|(e - 1)(a + \gamma)\| + \frac{1}{\beta + 1}|\gamma| \\ &\leqslant \|(e - 1)a\| + \left(\beta + 1 + \frac{1}{\beta + 1}\right)|\gamma| \\ &< t \quad \text{by (A) above.} \end{aligned}$$

Also, $z' - z = (a + \gamma)^{-1}(c(e) - 1)y$

$$= (a + \gamma)^{-1}\frac{1}{\beta}(e - 1)y ,$$

whence $\dfrac{(\beta + 1)^2}{\varepsilon}\|z' - z\| \leqslant \dfrac{(\beta + 1)^2}{\varepsilon}\|(a + \gamma)^{-1}\| \cdot \dfrac{1}{\beta}\|(e - 1)y\|$

$$< \frac{(\beta + 1)^2}{\varepsilon} \cdot \frac{1}{\beta} \cdot \beta \cdot \frac{\varepsilon}{(\beta + 1)^2} \cdot t$$

(by (B) above),

i.e. $\dfrac{(\beta + 1)^2}{\varepsilon}\|z' - z\| < t$.

But, by definition, $\|x' - x\| = \max\left\{\|a' - a\| , \dfrac{(\beta + 1)^2}{\varepsilon}\|z' - z\|\right\}$,

so $\|x' - x\| < t$.

It is not difficult to prove (vi) :

$$\|a'\| \leqslant \|a\| + \|a' - a\| ,$$

i.e. $\|a'\| \leqslant \beta^2 + 2\beta + \varepsilon - (\beta + 1)t + t$, i.e. $\|a'\| \leqslant \beta^2 + 2\beta + \varepsilon - \beta t$.

To prove (vii) note that

$$a' + \frac{\beta}{\beta + 1}\gamma = [c(e)]^{-1}(a + \gamma) ,$$

the product of two invertible elements.

Finally (viii) is immediate since

$$\left(a' + \frac{\beta}{\beta + 1}\gamma\right)z' = [c(e)]^{-1}(a + \gamma)(a + \gamma)^{-1}c(e).y$$
$$= y .$$

A further requirement before application of the Induction Theorem can be made is that at least one member of the family $Z(\cdot)$ is non-void. This is not difficult to show. There exists a number t_0 such that

$$\beta - \frac{\beta}{\beta + 1} \leqslant \beta^2 + 2\beta + \varepsilon - (\beta + 1)t_0 ,$$

and $$\frac{\beta}{\beta + 1} < \frac{\beta + 1}{\beta^2 + 2\beta + 2} \cdot t_0 .$$

These inequalities follow directly from the easily verified fact that

$$\beta^2 + 2\beta + \varepsilon > \beta - \frac{\beta}{\beta + 1} + (\beta^2 + 2\beta + 2)\frac{\beta}{\beta + 1} .$$

It now follows from Lemma 6.3.3 and the definition of $Z(\cdot)$ that if we set

$$a_0 = [c(e)]^{-1} - \frac{\beta}{\beta + 1} \quad \text{and} \quad z_0 = c(e)y ,$$

and taking γ as $\frac{\beta}{\beta + 1}$,

then $x_0 = (a_0, z_0) \in Z(t_0)$.

Now it is possible to apply the Induction Theorem (2.3.2). We have

$$\omega(t) = \frac{\beta}{\beta + 1} t ,$$

so that $\sigma(t) = (\beta + 1)t$.

Since $x_0 \in Z(t_0)$ there exists $x \in B(x_0, (\beta + 1)t_0) \cap Z(0)$. If $x = (a,z)$, we have $y = az$ and $\|a\| \leqslant \beta^2 + 2\beta + \varepsilon$. Furthermore, by definition of $\|x - x_0\|$,

$$\|z - y\| \leqslant \|z - z_0\| + \|z_0 - y\| \leqslant \frac{\varepsilon}{(\beta + 1)^2} \|x - x_0\| + \|c(e)y - y\| .$$

So, if we choose e in such a way that $\|c(e)y - y\| < \frac{\varepsilon}{(\beta + 1)^2}$,

we obtain $\|z - y\| < \frac{\varepsilon}{(\beta + 1)^2}(\beta + 1)t_0 + \frac{\varepsilon}{(\beta + 1)^2}$.

But $$(\beta + 1)t_0 \leqslant \beta^2 + 2\beta + \varepsilon - \beta + \frac{\beta}{\beta + 1} ,$$

i.e. $$(\beta + 1)t_0 \leqslant \beta^2 + \beta + \frac{\beta}{\beta + 1} + \varepsilon ,$$

which implies $(\beta + 1)t_0 < \beta^2 + 2\beta$ if we impose the restriction $\varepsilon < \frac{\beta^2}{\beta + 1}$.

So $\|z - y\| < \frac{\varepsilon}{(\beta + 1)^2} (\beta^2 + 2\beta + 1) = \varepsilon$.

6.4 Strict irreducibility of *-representations of B*-algebras

Topologically irreducible modules do not fit well into the general theory of representation of Banach algebras, but Kadison [6] proved the remarkable result that topologically irreducible Hilbert modules over B*-algebras are strictly irreducible. As our final example of the applications of the Induction Theorem we shall develop an unconventional proof of this result. Before doing so we require background material in the form of definitions, a variation of Schur's Lemma, a result concerning spectra in B*-algebras, another involving ideals in B*-algebras, and finally, a theorem by Kaplansky.

6.4.1 Background material

Let X be a normed linear space. We shall use the following notation:

$L(X)$ = the linear space of all linear mappings: $X \to X$;

$BL(X)$ = the linear space of all bounded linear mappings: $X \to X$.

A will denote an algebra over $\mathbb{R}$ or $\mathbb{C}$ and we shall use the concept of a left A-module as defined in 6.3.1.

6.4.1.1 Definitions If M is a linear space over $\mathbb{R}$ or $\mathbb{C}$, by a *representation of* A *on* M we shall mean a homomorphism of A into $L(M)$. Given a representation π of A on M, the *corresponding left* A*-module* is the linear space M with the module multiplication given by

(1) $am = \pi(a)m$ for each $a \in A$, $m \in M$.

M is called *non-trivial* if $AM \neq \{0\}$. The *corresponding representation of* A *on* M is the homomorphism π of A into $L(M)$ defined by (1) above. M is said to be *irreducible* if M and $\{0\}$ are its only A-submodules. A representation of A is called *irreducible* if the corresponding left A-module is irreducible.

M is said to be *normed* if it satisfies the axiom:

there exists a positive constant K such that

$$\|am\| \leqslant K\|a\|\|m\| \quad \text{for each} \quad a \in A,\ m \in M .$$

M is called a *Banach left* A-*module* if it is complete as a normed linear space. If A is a Banach algebra over $\mathbb{R}$ or $\mathbb{C}$ and M is a Banach left A-module we say that M is *topologically irreducible* if $A \neq \{0\}$, and $\{0\}$ and M are the only closed submodules of M. A *Hilbert* A-*module* is a Hilbert space $(H, \langle\ ,\ \rangle)$ which is a Banach left A-module such that

$$\langle ax, y\rangle = \langle x, a^*y\rangle \quad \text{for all} \quad x,y \in H , \quad a \in A .$$

Now let $(A,*)$ and $(B,\#)$ be star algebras. A *star homomorphism* is a homomorphism ϕ of A into B such that $\phi(a^*) = (\phi(a))^{\#}$ for each $a \in A$. Lastly, if $(H, \langle\ ,\ \rangle)$ is a Hilbert space, a *star representation* of A on H is a star homomorphism of A into $BL(H, \langle\ ,\ \rangle)$.

We shall make use of a topological variant of Schur's Lemma.

6.4.1.2 Lemma Let H be a Hilbert space and A any $*$-subalgebra of $B(H)$. Then, in order for A to be irreducible on H, it is necessary and sufficient that the centralizer A' of A in $B(H)$ reduces to scalar multiples of the identity operator.

A proof is to be found in Rickart [17], Chapter IV. We shall also use two more results from the same source.

6.4.1.3 **Theorem** If B is a Banach $*$-algebra which contains A as a $*$-subalgebra, then for every $x \in A$

$$\text{the spectrum of } x \text{ in } A \cup \{0\} = \text{the spectrum of } x \text{ in } B \cup \{0\} ,$$

If B is also a B^*-algebra, then the embedding of A in B is an isometry, so that A is a closed subalgebra of B .

6.4.1.4 **Theorem** Let I be a closed 2-sided ideal in A , a complex B^*-algebra. Then $I^* = I$ and A/I is also a B^*-algebra.

Our last requirement is a theorem by Kaplansky [7] concerning density in $*$-algebras of operators. This is also stated without proof.

6.4.1.5 **Theorem** Let M and N be $*$-algebras of operators on a Hilbert space with $M \subset N$ and M strongly dense in N . Then the unit sphere of M is strongly dense in the unit sphere of N .

Now we are ready to develop a proof of Kadison's Theorem *via* Pták's Induction Theorem [10].

6.4.2 **Theorem** A $*$-representation of a B^*-algebra on a Hilbert space is strictly irreducible if and only if it is topologically irreducible.

Proof Let $a' \to T_{a'}$ be any $*$-representation of a B^*-algebra A' on a Hilbert space H and denote by A the image of A' in $BL(H)$. Since $a' \to T_{a'}$ is automatically continuous, its kernel K is a closed 2-sided ideal in A'. Therefore, by Theorem 6.4.1.4, A'/K is also a B^*-algebra. Since A'/K is $*$-isomorphic with A, we conclude from Theorem 6.4.1.3 that A is a closed subalgebra of $BL(H)$. Therefore the problem reduces to showing that any C^*-algebra A which is topologically irreducible on its Hilbert space H is necessarily strictly irreducible.

If A is topologically irreducible on H, then, by Schur's Lemma (6.4.1.2), the centralizer of A in $BL(H)$ reduces to scalar multiples of the identity operator. Therefore the double centralizer coincides with $BL(H)$. It follows (see Dixmier [3]) that A is dense in $BL(H)$ relative to the strong neighbourhood topology for operators. Therefore, by Kaplansky's theorem (6.4.1.5), we have

(1) the unit ball of A is dense in the unit ball of $BL(H)$ with respect to the strong neighbourhood topology.

Now let x and x_0 be arbitrary elements of H with $x_0 \neq 0$. We are required to show the existence of $T \in A$ such that $Tx_0 = x$ and $\|T\| \leqslant (1+\varepsilon)\|x\|$.

If $x = 0$ it suffices to take $T = 0$. Hence we may suppose $x \neq 0$. Set $u = \frac{x}{\|x\|}$.

Then clearly we must find $U \in A$ such that

$$Ux_0 = u \quad \text{and} \quad \|U\| \leqslant 1 + \varepsilon$$

and set $T = \|x\| U$. It follows that we may restrict ourselves to x such that $\|x\| = 1$.

We define an approximate set $Z(\cdot)$ in A as follows: for each positive $t < 1$ let

$$Z(t) = \{T \in A : \|T\| \leqslant (1+\varepsilon)(1-t), \|Tx_0 - x\| < t\} .$$

We shall prove the following implication:

$$\text{if} \quad T \in Z(t) \quad \text{then} \quad B(T,t) \cap Z\left(\frac{\varepsilon}{1+\varepsilon}\, t\right) \neq \emptyset .$$

Suppose that $T \in Z(t)$. Let V_0 be the one-dimensional operator defined by

$$V_0 y = (y,x_0)(x - Tx_0) .$$

Then $\|V_0\| = \|x - Tx_0\| < t$. And, according to (1), there exists $V \in A$ such that $\|V\| \leqslant \|V_0\|$ and

$$\|Vx_0 - (x - Tx_0)\| = \|(V - V_0)x_0\| < \frac{\varepsilon}{1+\varepsilon}\, t .$$

Now let $T' = T + V$. We have

$$\|T' - T\| = \|V\| \leqslant \|V_0\| < t ,$$

$$\|T'x_0 - x\| < \frac{\varepsilon}{1+\varepsilon}\, t ,$$

and $\|T'\| \leqslant \|T\| + \|T' - T\| \leqslant (1+\varepsilon)(1-t) + t$,

i.e. $\|T'\| \leqslant (1+\varepsilon)\left(1 - t + \frac{1}{1+\varepsilon}\, t\right)$,

i.e. $\|T'\| \leqslant (1+\varepsilon)\left(1 - \frac{\varepsilon}{1+\varepsilon}\, t\right)$.

It follows that $T' \in B(T,t) \cap Z\left(\frac{\varepsilon}{1+\varepsilon}\, t\right)$.

Now, to complete the proof, we show that at least one $Z(t)$ is nonvoid. Consider the operator T_0 defined by $T_0 y = (y,x_0)x$. We have $T_0 x_0 = x$ and

$$\|T_0\| = \|x_0\| . \|x\| = 1 = (1+\varepsilon)\left(1 - \frac{\varepsilon}{1+\varepsilon}\right) .$$

Thus $T_0 \in Z\left(\frac{\varepsilon}{1+\varepsilon}\right)$. And, using the Kaplansky Density

Theorem (6.4.1.5), there exists $T_1 \in A$ such that $\|T_1\| \leq \|T_0\|$ and $\|T_1x_0 - T_0x_0\| < \frac{\varepsilon}{1+\varepsilon}$. So $T_1 \in Z\left(\frac{\varepsilon}{1+\varepsilon}\right)$ and there exists t such that $Z(t) \neq \emptyset$.

Now we may apply the Induction Theorem (2.3.1), the functions q and p being defined by

$$q(t) = t \quad \text{and} \quad p(t) = \frac{\varepsilon}{1+\varepsilon}t \ ;$$

hence $$\omega(t) = \frac{\varepsilon}{1+\varepsilon}t \quad \text{and} \quad \sigma(t) = (1+\varepsilon)t \ .$$

By the Induction Theorem $Z(0) \neq \emptyset$ and, of course, any element of $Z(0)$ is a solution.

REFERENCES

[1] F.F. Bonsall and J. Duncan *Complete Normed Algebras*, Springer-Verlag, Berlin, 1973.

[2] J. Dieudonné *Foundations of Modern Analysis*, Academic Press, New York and London, 1963.

[3] J. Dixmier Les algèbres d'opérateurs dans l'espace hilbertien (Algèbres de von Neumann), *Cahiers Scientifique*, Fas. XXV, Gauthier-Villars, Paris, 1957.

[4] M. Edelstein An extension of Banach's Contraction Principle, *Proc. Amer. Math. Soc.* 12 (1961), 7 - 10.

[5] M. Fiedler and V. Pták Some estimates and iteration procedures for the spectrum of an almost decomposable matrix, *Czech. Math. J.* 89 (1964), 593-608.

[6] R.V. Kadison Irreducible operator algebras, *Proc. Nat. Ac. of Sci. U.S.A.* 43 (1957), 273-276.

[7] I. Kaplansky A theorem on rings of operators, *Pacific J. Math.* 1 (1951), 227-232.

[8] J.L. Kelley *General Topology*, D. Van Nostrand Co. Inc., Princeton, New Jersey, 1964.

[9] I. Niven and H.S. Zuckerman *An Introduction to the Theory of Numbers*, 3rd ed., John Wiley and Sons Inc., New York, 1972.

[10] V. Pták A theorem of the closed graph type, *Man. Math.* 13 (1974), 109-130.

[11] V. Pták A quantitative refinement of the closed graph theorem, *Czech. Math. J.* 24 (99) (1974), 503-506.

[12] V. Pták Nondiscrete mathematical induction and iterative existence proofs, *Linear Algebra and its Applications,* 13 (1976), 223-238, Amer. Elsevier Publ. Co. Inc., 1976.

[13] V. Pták The rate of convergence of Newton's Process, Prepublication Copy, Prague, April 1975.

[14] V. Pták Concerning the rate of convergence of Newton's Process, *Comm. Math. Univ. Car.,* 16 (1975), 699-705.

[15] V. Pták A rate of convergence, Prepublication Copy, Prague, Undated.

[16] V. Pták A modification of Newton's Method, *Čas. pěst. mat.,* 101 (1976), 188-194.

[17] G.F. Simmons *Introduction to Topology and Modern Analysis,* McGraw-Hill, Tokyo, 1963.

[18] N.E. Steenrod, P.R. Halmos, M.M. Schiffer and J.A. Dieudonné *How To Write Mathematics*, American Mathematical Society, 1975.

[19] A.E. Taylor *Introduction to Functional Analysis,* John Wiley and Sons Inc., New York, 1958.