



UNIVERSIDADE DO MINHO

ESCOLA DE CIÊNCIAS

Raul Aparício Gonçalves

**Utilização de métodos numéricos na
resolução de equações e perspetivas de
integração curricular no Ensino
Secundário**

Tese de Mestrado

Mestrado em Ciências – Formação Contínua de Professores,
Área de Especialização em Matemática

Trabalho efetuado sob a orientação do

Professor Rui Manuel da Silva Ralha

outubro de 2014

DECLARAÇÃO

Raul Aparício Gonçalves

Endereço Eletrónico: prof.raulaparicio@gmail.com

Telefones: 224224642 / 966657429

Cartão de Cidadão nº 8377503 0 ZZ0

Tese de Mestrado: Utilização de métodos numéricos na resolução de equações e perspectivas de integração curricular no Ensino Secundário

Orientador: Rui Manuel da Silva Ralha

Ano da Conclusão: 2014

Mestrado em Ciências – Formação Contínua de Professores, Área de Especialização em Matemática

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE

Universidade do Minho, 31 de Outubro de 2014

Assinatura: _____

“(...) se alguém lhes perguntar como se calculam todas as raízes de uma dada equação algébrica, de grau arbitrário, com a aproximação que se queira, terão de reconhecer que não sabem. Isto dá bem nota de como o ensino tradicional tem sido afastado da realidade.”

José Sebastião e Silva

AGRADECIMENTOS

A todos aqueles que contribuíram de forma mais ou menos decisiva para que este trabalho se tivesse realizado e aos que futuramente contribuirão para que ele possa ser útil.

Em particular aos Professores Rui Ralha e Elfrida Ralha, que numa fase inicial me deram incentivo e me convenceram de que eu seria capaz de realizar um trabalho de dissertação que integrasse os meus interesses, quer ao nível de conteúdo quer ao nível do cumprimento de prazos e sem que as alterações da gestão do tempo diário causassem prejuízo para os meus afazeres profissionais.

Em especial ao professor Rui Ralha, que ao longo do ano foi acompanhando o meu trabalho na qualidade de orientador, sobretudo pela forma simples com que me ajudou a estudar e a compreender alguns assuntos complexos e também pela forma como apresentou as suas sugestões de correção aos documentos que ia escrevendo.

À minha família, em especial à Sónia, ao Rui Pedro e à Inês.

UTILIZAÇÃO DE MÉTODOS NUMÉRICOS NA RESOLUÇÃO DE EQUAÇÕES E PERSPETIVAS DE INTEGRAÇÃO CURRICULAR NO ENSINO SECUNDÁRIO

RESUMO

Esta dissertação realiza-se no âmbito do Mestrado em Ciências (Formação Contínua de Professores) cujo público-alvo são sobretudo os professores do Ensino Secundário que desejem aprofundar conhecimentos e desenvolver competências com vista a uma melhoria do seu desempenho como professores.

Em função da experiência profissional do autor e da pesquisa bibliográfica realizada percebeu-se que o trabalho com o cálculo numérico pode ter um grande potencial na formação matemática dos alunos nos anos que antecedem a entrada num curso superior. Parece-nos de primordial importância confrontar os alunos em fase final do Ensino Secundário com problemas que não podem ser resolvidos por fórmulas exatas e sensibilizá-los para a importância dos métodos de cálculo aproximado. O tema de equações não lineares adequa-se completamente a este objetivo. Nos programas tradicionais os alunos aprendem a usar a fórmula resolvente das equações polinomiais do 2º grau e também a regra de Ruffini se uma das raízes for de determinação fácil. Depois de aprendidas estas regras clássicas, devem os alunos ser confrontados com outras equações (polinomiais ou não) que mostrem a necessidade de outras ferramentas matemáticas e computacionais.

Assim, este trabalho é uma reflexão sobre esta temática e inclui a apresentação de exemplos concretos da possibilidade de trabalho destas questões com alunos do Ensino Secundário. Do vasto leque de métodos numéricos que existem para a resolução de equações selecionamos apenas três (dos mais populares) que servem bem o objetivo de perspetivar a respetiva integração curricular no ensino não superior. São eles o método da bisseção, o método do ponto fixo e o método de Newton-Raphson. Para além do respetivo interesse intrínseco, estes métodos são

também uma oportunidade para os alunos apreciarem uma aplicação muito concreta de conhecimentos matemáticos por eles já adquiridos (limites de sequências numéricas, o teorema de Bolzano, derivadas e o teorema do valor médio de Lagrange, entre outros). Para além disto, a implementação prática destes métodos também envolve capacidades que os alunos desenvolvem em unidades curriculares da área da informática.

Uma questão fundamental na utilização dos métodos numéricos é a do controle dos erros. Os erros de arredondamento são inevitáveis porque a máquina de calcular (ou o computador) opera apenas com um número finito, embora grande, de números e com eles tem que representar todos os números reais (dentro de certos limites). Por outro lado, os erros propagam-se de diferentes maneiras em diferentes sequências de operações numéricas. Estes são temas complexos tratados em cursos superiores e é nosso entendimento que não deve ser tentado no ensino antes disso. No entanto, pode ser interessante dar aos alunos alguma informação sobre a representação de números em ponto flutuante, quer no sistema decimal quer no sistema binário, e mostrar com exemplos algumas consequências na precisão da aritmética computacional.

THE USE OF NUMERICAL METHODS FOR SOLVING EQUATIONS AND PERSPECTIVES FOR CURRICULUM INTEGRATION IN HIGH SCHOOL

ABSTRACT

This dissertation takes place within an MSc degree in Science which is tailored for those which are already professional teachers in high school but envisage to acquire further knowledge and skills that will, hopefully, help them in their teaching duties. In the light of the author's teaching experience and the biographical research conducted, it was realized that working with numerical calculus may have great potential in mathematics education of students even before their entrance into higher education. In our opinion, it is of paramount importance to make the students, at a late stage of their high school learning, tackle problems that cannot be solved by exact formulae and raise their attention to the relevance of methods that provide numerical approximations. The topic of solving equations serves well such goal. In the traditional curricula students learn how to use the exact formula for the roots of quadratic equations and also Ruffini's rule if they manage somehow to get one of the roots in the first place. But after this the students should be given other equations (not necessarily polynomial) that make evident the need of further mathematical and computational tools.

Thus, this work aims to be a study on the subject and, in particular, we present examples of how these topics may be taught to high school students. From the large spectrum of numerical methods that exist for solving equations we selected only three (of the most popular) that are well in accordance with our goal of their inclusion in pre-university curricula. These are the bisection method, the fixed point iteration and Newton-Raphson's method. Besides their own importance, these methods are also an opportunity for the students to see a very concrete application of mathematical knowledge that they should have learned before the use of numerical methods (limits of sequences of numbers, Bolzano's theorem, derivatives and Lagrange's mean value theorem, among others). Furthermore, the practical implementation of the methods also require skills that the students are expected to develop in curriculum units that belong to the area of information technology.

A fundamental issue in using numerical methods is that of keeping the errors under control. Rounding errors are unavoidable because the calculator (or the computer) only stores a large but finite number of representations. Furthermore, the analysis of propagation of numerical errors belongs to advanced courses where the numerical stability of algorithms is studied and clearly no attempt should be made to treat such subject in high school. However, it may be useful to teach the students about the way numbers are represented in the calculator and in the computer, and show with examples some consequences on the precision of the computational arithmetic.

.

ÍNDICE

1 Introdução	001
2 Aritmética para máquinas de cálculo	005
2.1 Introdução	005
2.2 Representação de números em máquinas de cálculo	012
2.3 Modos de arredondamento	016
2.4 Erros numéricos	017
2.4.1 Erros de arredondamento	020
2.4.2 Propagação de erros nas operações aritméticas	021
3 Alguns métodos numéricos de resolução de equações	029
3.1 Introdução	029
3.2 Método da bisseção	031
3.2.1 Introdução	031
3.2.2 Funcionamento do método	031
3.2.3 Alguns constrangimentos	035
3.2.4 Breve conclusão	038
3.3 Método do ponto fixo	039
3.3.1 Ponto fixo	039
3.3.2 Ponto fixo e resolução de equações	039
3.3.3 Método iterativo baseado na aproximação ao ponto fixo	042
3.3.4 Condições de convergência do método do ponto fixo	045
3.3.5 O método do ponto fixo na resolução de equações do 2º grau	049
3.4 Método de Newton-Raphson	053
3.4.1 Introdução e nota histórica	053
3.4.2 Interpretação geométrica e algoritmo	054
3.4.3 O método de Newton-Raphson e a fórmula de Taylor	056
3.4.4 Alguns constrangimentos	059
3.4.5 Convergência do método	064
3.4.6 Conclusão	076

3.5	Comparação dos métodos estudados	077
3.5.1	Introdução	077
3.5.2	Alguns exemplos	078
3.5.3	Conclusão	085
3.6	O cálculo de raízes em máquinas de calcular	086
4	Conclusões e perspetivas de integração curricular	093
4.1	Introdução	093
4.2	Algumas considerações gerais	093
4.3	Tarefas para a sala de aula	098
	Bibliografia	109

1. INTRODUÇÃO

No meu percurso profissional de professor do Ensino Básico e Secundário, que já leva mais de 20 anos, a tecnologia no Ensino tem ocupado um lugar de destaque nas minhas ações. Enquanto estudante universitário, na licenciatura de Matemática – Ramo Educacional, que terminei no ano de 1992, na Faculdade de Ciências da Universidade do Porto, fascinei-me com a utilização do meu primeiro computador, um Timex Sinclair 2068 (figura 1.1) que, apesar de bastante mais limitado do que os atuais, me permitiu desenvolver capacidades de programação, sobretudo em BASIC. Nas aulas de Análise Numérica a computação era feita com uma máquina de calcular com funcionalidades de programação, a TI-66 Programable, (figura 1.2) sendo a linguagem de programação própria da empresa que a comercializou, mas com comandos muito idênticos aos do BASIC.



Fig. 1.1



Fig. 1.2

Da minha aprendizagem no âmbito desta cadeira ficaram sobretudo o gosto pela utilização das máquinas e a sensibilidade para a importância dos erros numéricos e dos métodos de cálculo aproximado de soluções. Como professor não tenho tido a oportunidade para desenvolver tanto quanto gostaria estas temáticas, que não têm feito parte dos conteúdos dos programas oficiais de matemática.

Já no exercício profissional desempenhou papel importante na minha formação a participação no “Programa de Acompanhamento dos Programas de Matemática do Ensino Secundário”, do Ministério da Educação, a cujo grupo de trabalho tive a oportunidade e o prazer de pertencer, entre 1998/1999 e 2001/2002. Os meus horizontes do ensino da matemática alargaram-se de forma sustentada com uma utilização mais intensiva das calculadoras gráficas e os desafios que isso acarretou, com mais conexões entre os diferentes temas de matemática e com a realização de atividades extracurriculares com os alunos sobre problemas matemáticos.

Num contexto de explorações para tornar as aprendizagens dos meus alunos do 12º ano, em 2001, mais significativas, surgiu a análise do comportamento relativo das funções definidas por x^2 e por 2^x [1]. Tal levou a uma tão grande curiosidade por parte de alguns alunos na obtenção da abcissa de um dos pontos de intersecção dos gráficos das duas funções que, só um trabalho sobre a forma como a calculadora gráfica funciona para solucionar equações poderia satisfazer essa curiosidade. O trabalho desenvolvido com os alunos para dar resposta a estas questões suscitou-me um conjunto de reflexões que levaram a algumas suposições e constatações, nomeadamente:

- um professor precisa de um aprofundamento científico que ultrapasse muito o que ensina aos alunos, em particular, a compreensão de certos resultados produzidos na calculadora gráfica (obrigatória desde os programas do Ensino Secundário de 1997 e de 2003, e prevista nos programas que se iniciarão no ano letivo de 2015/2016 a partir do 10º ano) requer o conhecimento da forma de funcionar da máquina e as suas limitações;
- alguns alunos são capazes de trabalhar com assuntos de nível científico superior ao que os currículos contemplam, pelo que há necessidade de elevar as expectativas em relação a tais alunos e levá-los a desenvolver pequenos projetos, individuais ou em grupo, onde se apliquem e consolidem os conhecimentos adquiridos nas aulas;
- o tratamento de alguns temas de nível superior, que sejam do agrado dos alunos, como a utilização de certos métodos numéricos, pode favorecer a compreensão de temas curriculares, quer durante o seu tratamento, quer após esse tratamento, como revisão e conexão entre partes menos ligadas, sendo a ligação com a tecnologia um elemento que favorece a apetência dos alunos por estas questões.

O tratamento de temas de análise numérica por alunos do Ensino Pré-universitário é um assunto que tem merecido a atenção de investigadores/educadores. William P. Fox e Richard D. West,

como pode observar-se em [2], têm longa experiência, nos Estados Unidos, na integração de métodos numéricos no Ensino Pré-universitário, no âmbito do que denominam de projetos interdisciplinares. Referem, por exemplo, que a compreensão das medidas do erro é um dos principais objetivos de qualquer curso introdutório de análise numérica e consideram que um algoritmo simples, como o da bisseção, proporciona uma primeira oportunidade para enfrentar a análise do erro. Usam muitas vezes estes projetos de modo sumativo, depois da aprendizagem inicial ter tido lugar, porque tudo isto ajuda os alunos a uma aprendizagem mais sólida.

Na Austrália, no Estado de Victoria, como se pode observar em [3], depois de 10 anos de escolaridade os alunos desenvolvem, durante 4 semestres, projetos que visam uma certa especialização de acordo com a carreira que pretendem seguir. No caso da continuação dos estudos em matemática, uma das áreas visa o trabalho com métodos numéricos, como é referido por David-Leigh e Michael Evans em [4]. Estes autores advogam a utilização de máquinas ou computadores para resolução numérica de equações, derivação, integração, bem como a utilização do método de Euler para equações diferenciais simples.

Também em Portugal aparecem algumas referências à utilização de métodos numéricos no currículo do Ensino Secundário, umas vezes concretizadas na prática, como situações pontuais em cursos não necessariamente de matemática [5], outras vezes perspetivando a vantagem da concretização num futuro próximo [6]. Num período em que os computadores e máquinas de calcular eram ainda muito pouco usados, Sebastião e Silva [6] já defendia que estes meios deveriam desempenhar um papel fundamental no ensino da matemática, provavelmente perspetivando a generalização do seu uso num futuro não longínquo. Sebastião e Silva escreve que devem os alunos ser postos em contacto com os métodos de análise numérica “moderna”, isto é, ligada aos computadores. Como exemplo do afastamento entre o ensino “tradicional” e a vida real, Sebastião e Silva critica o facto de aos alunos não serem ensinados processos que lhes permitam calcular as raízes de uma equação algébrica de grau arbitrário. Esta crítica particular é dirigida pelo excelente professor e pedagogo ao Ensino Superior de então e a nossa opinião é a de que, no momento presente, se aplica à realidade do Ensino Secundário. Sebastião e Silva também observou, por exemplo, que os processos de recorrência, baseados no princípio da indução matemática, constituíam um dos muitos assuntos da matemática postos na ordem do dia pelos computadores.

O desejo do autor é que esta dissertação possa ser de alguma utilidade também para outros professores e que sobre estes temas reflitam e partilhem ideias e experiências entre si. Para este

fim, procurámos escrever um texto de leitura acessível e que desperte o interesse do “professor leitor”. Assim, por um lado, a “teoria”, embora rigorosa do ponto de vista matemático, é reduzida ao estritamente necessário para um bom entendimento do funcionamento dos métodos numéricos; por outro lado, são trabalhados numerosos exemplos, com apresentação de resultados obtidos em algumas calculadoras gráficas e o leitor interessado poderá, munido da sua própria calculadora, repetir a experiência relatada. O autor espera que a sua já longa experiência como professor lhe tenha permitido encontrar uma “linguagem” adequada para a comunicação com outros professores.

Este trabalho contempla quatro capítulos devidamente ligados e justificados em conjunto. Optou-se pelo tratamento de três métodos numéricos, o método da bissecção, o método do ponto fixo e o método de Newton-Raphson. Este último, que tem uma longa história, é ainda hoje um dos mais importantes métodos iterativos para a determinação de raízes de equações (e não só). Além disso, consideramos que o seu estudo, pelo menos nos aspetos mais básicos, está ao alcance dos alunos em fase final do Ensino Secundário e permite fazer a ligação com o tema das derivadas. O método da bissecção é um método muito simples e robusto (tem convergência garantida) e relaciona-se com o teorema de Bolzano, o qual consta no programa oficial do Ensino Secundário. O método do ponto fixo tem também grande potencial formativo ao nível do Ensino Básico e Secundário, sendo um método que associa normalmente uma maior rapidez em relação ao método da bissecção e a ultrapassagem de algumas dificuldades que possam decorrer da utilização do método de Newton-Raphson. Para compreender alguns aspetos do funcionamento dos métodos, e porque estes lidam com sucessões e valores aproximados, importa a análise do erro e da representação e números na máquina, o que é feito já a seguir, após esta introdução.

2. ARITMÉTICA PARA MÁQUINAS DE CÁLCULO

2.1. INTRODUÇÃO

A procura de formas de facilitar o cálculo ocupa a humanidade há milénios. Surge no séc. XVII a primeira calculadora mecânica pela mão de Wilhelm Schickard (1592-1635), um pouco antes da conhecida pascalina (figura 2.1) de Blaise Pascal (1623-1662), por ele construída com o objetivo de ajudar o pai, que trabalhava nos impostos. Schickard e Pascal são considerados os precursores das máquinas de calcular mecânicas, as quais foram sofrendo evoluções que acompanharam os avanços da tecnologia mecânica até ao séc. XX. Note-se que apenas em 1810 viria a ser construída a primeira máquina de calcular comercial. No final do séc. XIX as máquinas comerciais generalizaram-se, sendo famosa a máquina que o estatístico americano Herman Hollerith (1860-1929), principal impulsionador do leitor de cartões perfurados e um dos fundadores da IBM, construiu para tratar os dados do censo de 1890 nos Estados Unidos (figura 2.2) [7], conhecida por Tabuladora de Hollerith.



Fig. 2.1.

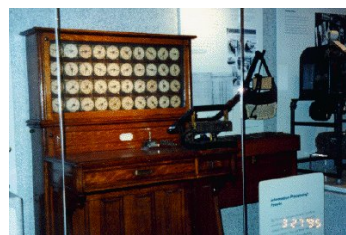


Fig. 2.2. [8]

A última calculadora mecânica terá sido, porventura, a Curta (figura 2.3), uma calculadora mecânica de bolso que apareceu em 1948 pela mão do engenheiro austríaco Curt Herzstark e que foi produzida na cidade de Vaduz no Principado de Liechtenstein.



Fig. 2.3

Para além das quatro operações elementares, a Curta permitia também o cálculo de raízes com diferentes índices, entre outras operações.

Uns dos procedimentos para efetuar o cálculo da raiz quadrada numa Curta, que pode ser observado em documentos originais publicados em [9] consistia em partir de uma estimativa inicial R da raiz quadrada N e calcular a soma dos quadrados de N e de R e depois multiplicar o resultado pelo inverso do dobro de R . Nos capítulos seguintes será novamente referida esta forma de aproximar a raiz quadrada de um número num contexto de estudo de métodos numéricos de resolução de equações.

A partir de 1930, o desenvolvimento das diversas ciências e a sua aplicação a setores cada vez mais numerosos da atividade humana exigiam cálculos cada vez mais complexos, a que o aparecimento dos primeiros computadores baseados em dispositivos eletrónicos, por altura da Segunda Guerra Mundial, veio dar resposta [7]. Desde então a evolução tem sido muito rápida e a primeira calculadora eletrónica de bolso surge no início da década de 70, a Pocketronic (figura 2.4), lançada pela Cannon em resultado de um trabalho prévio de desenvolvimento com a Texas Instruments [10].

A primeira calculadora científica de bolso, a HP-35 [11], surgiu em 1972. Esta calculadora eletrónica, cujo nome se deve ao número de teclas que possui, trouxe pela primeira vez operações e funcionalidades como as funções trigonométricas, logarítmicas, exponenciais e outras funções transcendentais, para além das quatro operações usuais (figura 2.5).

Em 1985 aparece a 1ª calculadora científica com funcionalidade gráfica, lançada pela CASIO, a fx-7000G (figura 2.6).



Fig. 2.4



Fig. 2.5

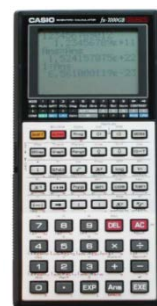


Fig. 2.6

São vários os modelos comercializados atualmente, escolhendo-se três dos mais utilizados pelos estudantes, sobretudo do Ensino Secundário, para figurarem neste trabalho. São eles o modelo CG-20 (figura 2.7), da Casio, e os modelos TI-84 Plus C SE (figura 2.8) e TI-Nspire Cx, (figura 2.9), da Texas Instrument. Alguns modelos incluem ainda o Cálculo Algébrico Simbólico (CAS), como é o caso da TI-Nspire Cx CAS (figura 2.10), por muitos considerado o modelo tecnologicamente

mais avançado alguma vez produzido. Nos últimos anos tem sido possível utilizar as calculadoras gráficas no ecrã do computador através de emuladores, favorecendo a exposição em aula ou a edição de ecrãs em documentos, sendo aqui a tecnologia TI-Nspire a mais avançada, na medida em que o software de computador vai muito mais além do que a repetição da calculadora no computador. A calculadora TI-Nspire é conhecida por “unidade portátil” (“handheld”), o que traduz a ideia de apêndice físico e interligado a um software com imensas valências, entre as quais a possibilidade de realização de páginas interativas para colocação na web ou um sistema de comunicação a distância (via wireless) com controle das unidades portáteis e análise de respostas a questionários (TI-Nspire Navigator).



Fig. 2.7

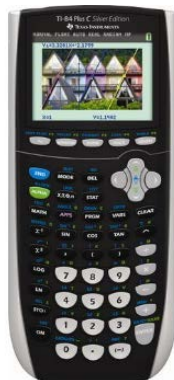


Fig. 2.8

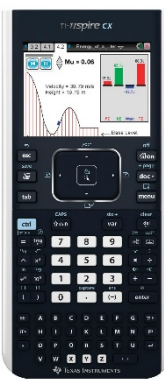


Fig. 2.9



Fig. 2.10

Só para dar uma ideia do avanço que ocorreu na miniaturização dos circuitos integrados, que possibilitou o aparecimento da calculadora eletrónica portátil, é referido em [12] que tal levou a uma redução do custo das funções eletrónicas de um milhão para um, evolução nunca conseguida antes com tanta magnitude. As calculadoras atuais trouxeram incrível facilidade aos cálculos. Quando pretendemos efetuar cálculos com números reais sem uma máquina, como uma calculadora ou um computador, habitualmente deparamo-nos com dificuldades que esses instrumentos nos permitem ultrapassar. No entanto, as suas limitações físicas, variáveis de máquina para máquina, obrigam a uma aritmética diferente da que utilizamos em cálculo ou álgebra.

Vejamos os resultados produzidos por três calculadoras gráficas, das mais utilizadas atualmente pelos alunos do Ensino Secundário, em alguns cálculos. Na figura 2.11 apresentam-se os resultados obtidos nas calculadoras indicadas para

$$987654321 + 0.0004 ,$$

$$987654321 + 0.0001$$

e

$$(987654321 + 0.0004) - (987654321 + 0.0001)$$

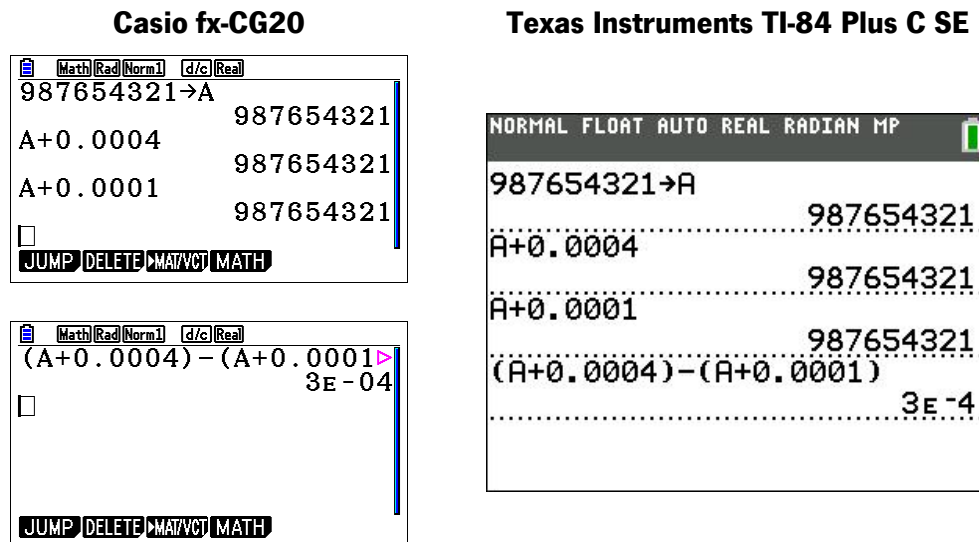


Fig. 2.11

Acrescentamos agora mais um algarismo nos números anteriores e calculamos

$$9876543210 + 0.0004,$$

$$9876543210 + 0.0001$$

e

$$(9876543210 + 0.0004) - (9876543210 + 0.0001)$$

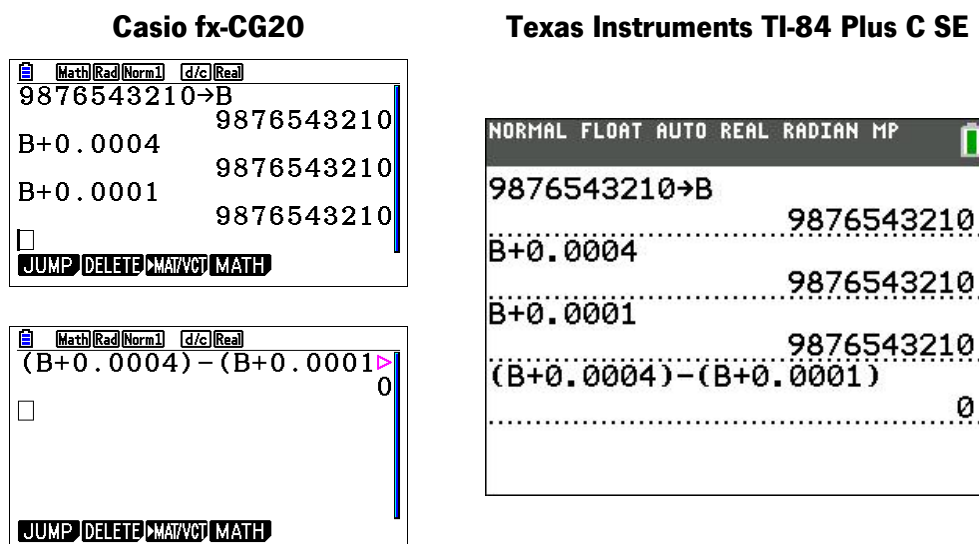


Fig. 2.12

Na figura 2.11 podemos observar que embora as somas $987654321+0.0004$ e $987654321+0.0001$ não produzam o mesmo valor (já que a diferença entre elas calculada pelas máquinas não é nula), as representações que aparecem no ecrã coincidem. Esta prática de esconder alguns dos últimos algarismos que fazem parte da representação interna de cada número é comum a praticamente todas as máquinas. Há portanto que ter presente que o número

que se vê no ecrã não é necessariamente o número que está representado internamente na máquina.

Na figura 2.12 a situação é diferente: as somas $9876543210+0.0004$ e $9876543210+0.0001$ têm a mesma representação interna e, por isso, a diferença é nula.

Dos dois exemplos anteriores, concluímos que as máquinas de calcular Casio fx-CG20 e TI-84 Plus C SE, têm capacidade para representar exatamente números com treze dígitos mas não com catorze dígitos. O facto de estas máquinas não representarem exatamente mais do que um determinado número de dígitos é comum a todas as máquinas, embora esse número máximo de dígitos possa variar de máquina para máquina. No modelo Texas Instruments Nspire CX é possível representar de forma exata mais um dígito que nos modelos anteriores.

Com efeito, como se pode constatar na figura 2.13, a diferença $x - y$ é exata com os números

$$x = 9876543210 + 0.0004$$

$$y = 9876543210 + 0.0001$$

que têm 14 algarismos, mas já não é exata com os números

$$x = 98765432100 + 0.0004$$

$$y = 98765432100 + 0.0001$$

que têm um algarismo mais.

Texas Instruments TI-Nspire CX

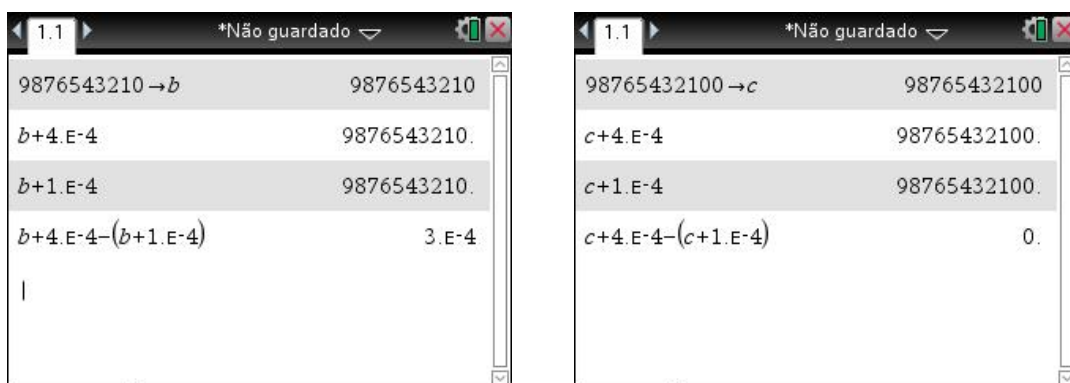


Fig. 2.13

Um “truque” útil para revelar algarismos que constam da representação interna de um número mas que não aparecem no ecrã da máquina é o de subtrair a esse número outro que resulta do primeiro por corte de alguns dos últimos algarismos. Na tabela 2.1, apresentam-se os resultados deste procedimento com as representações dos números:

$$\frac{13}{7} = 1.8571428571428571428571 \dots,$$

$$\frac{16}{7} = 2.285714285714285714285 \dots$$

e

$$\frac{18}{7} = 2.5714285714285714285714 \dots$$

Para cada um destes números, a Casio e a TI-84 apresentam apenas dez algarismos enquanto que a TI-Nspire apresenta doze algarismos. A subtração da representação com os primeiros dez algarismos permite revelar um total de catorze algarismos no caso de ambos os modelos da TI e de quinze algarismos no caso do modelo Casio.

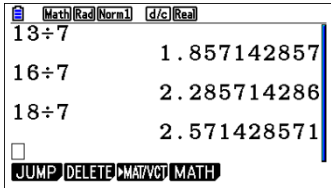
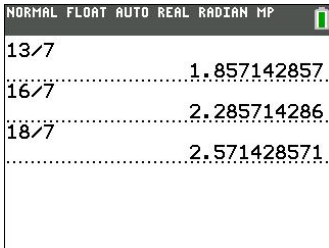

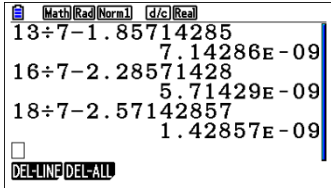
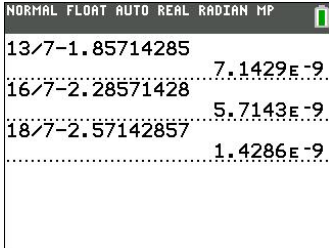

	Casio fx-CG20		TI-84 Plus C SE		TI-Nspire CX	
$\frac{13}{7}$	1.857142857	10	1.857142857	10	1.85714285714	12
	1.85714285714286	15	1.8571428571429	14	1.8571428571429	14
$\frac{16}{7}$	2.285714286	10	2.285714286	10	2.28571428571	12
	2.28571428571429	15	2.2857142857143	14	2.2857142857143	14
$\frac{18}{7}$	2.571428571	10	2.571428571	10	2.57142857143	12
	2.57142857142857	15	2.5714285714286	14	2.5714285714286	14
						
						

Tabela 2.1

Observa-se que a Casio fx-CG20 é a que “revela” mais dígitos sem que com isso seja mais rigorosa nos cálculos como se pode concluir do exemplo ilustrado nas figuras 2.12 e 2.13 (como se viu, ao contrário da Casio fx-CG20, a TI-Nspire CX produz o resultado exato).

Observa-se também que o último dígito observável é o próprio se o seguinte da representação exata é inferior a 5 e é superior em uma unidade no caso contrário, isto é, é usado um

arredondamento simétrico [13, p. 22]. É caso para dizer que “quem vê caras não vê corações”, para utilizar uma expressão popular.

De qualquer modo, observamos exemplos que permitem afirmar que há operações ao alcance de um aluno do 1º Ciclo do Ensino Básico que não se conseguem fazer exatamente numa calculadora gráfica. Isto não diminui a importância destes potentes instrumentos, em que uma das valências é o cálculo, mas sugere-nos a importância e a necessidade de compreender o seu funcionamento e limitações para uma utilização crítica e eficaz, sobretudo quando os limites são atingidos.

As limitações físicas das máquinas de cálculo levam a que não se possam representar todos os números reais. O conjunto \mathbb{R} é infinito, ilimitado e contínuo. Os números que se podem representar numa máquina têm de pertencer necessariamente a um conjunto finito, discreto e limitado, subconjunto de \mathbb{Q} . Denotaremos tal conjunto por \mathcal{F} .

Deste modo, quando se pretende utilizar numa máquina um número real x que não pertença a \mathcal{F} , ou seja, um número que não se pode representar exatamente na máquina, será substituído por um valor de \mathcal{F} , próximo de x , que denotaremos por $fl(x)$.

A figura 2.14 ilustra a problemática da representação de um número real $x \notin \mathcal{F}$ por um elemento de \mathcal{F} . Neste caso, $fl(x)$. Note-se que sendo \mathcal{F} um conjunto discreto, x ($x \notin \mathcal{F}$) pode enquadrar-se entre dois números consecutivos de \mathcal{F} , um **antecessor** (x_-) e um **sucessor** (x_+), de tal modo que $x_- < x < x_+ \wedge \nexists f \in \mathcal{F}: x \neq f \wedge x_- < f < x_+$, desde que x , x_- e x_+ não excedam os limites do conjunto \mathcal{F} , uma vez que é limitado.

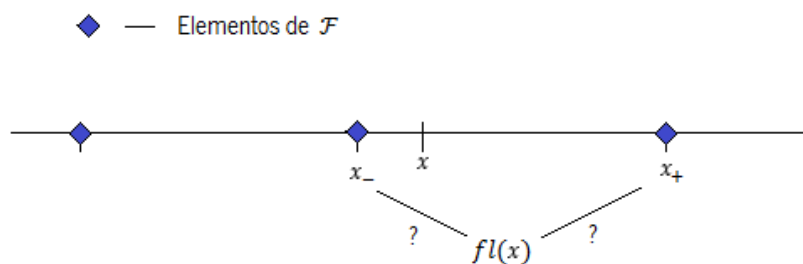


Fig. 2.14

$fl(x) = x$, $fl(x) = x_-$ ou $fl(x) = x_+$?

Quando o número real x é um elemento de \mathcal{F} , tem-se que $fl(x) = x$. Quando tal não acontece, a escolha entre x_- e x_+ depende do modo de arredondamento usado pela máquina e também das distâncias $x - x_-$ e $x_+ - x$.

Usualmente, as máquinas de calcular usam o arredondamento simétrico, mas existem outros modos de **arredondamento**, que desenvolveremos um pouco mais à frente.

2.2. REPRESENTAÇÃO DE NÚMEROS EM MÁQUINAS DE CÁLCULO

Para melhor compreender a temática do arredondamento é necessário fazer mais do que uma breve referência e para tal é necessária uma análise mais cuidada ao sistema de representação de números numa máquina, denominado por **sistema de ponto flutuante**. Neste sistema, cada número é representado na forma $\pm m \times b^e$ onde b , inteiro não inferior a 2, é a base do sistema, e o expoente (inteiro, positivo ou negativo), e m ($1 \leq m < b$) o conjunto dos algarismos significativos do número a que se dá usualmente o nome de mantissa. O conjunto \mathcal{F} pode assim variar de máquina para máquina, dependendo de três números inteiros: a base b , o número p de algarismos da mantissa e o número q de algarismos do expoente. Por esta razão, é costume escrever-se $\mathcal{F}(b, p, q)$, com valores b , p e q fixados, para representar o conjunto dos números de ponto flutuante desse sistema.

As calculadoras usam normalmente o sistema decimal, isto é, $b = 10$. Das experiências realizadas antes, conjecturamos ser $p = 13$ no caso da Casio fx-CG 20 e da TI-84 Plus C SE e $p = 14$ no caso da TI-Nspire CX. Para a Casio e para a TI-84, tem-se a gama de expoentes $-99 \leq e \leq 99$ (isto é, $q = 2$) enquanto que no caso da TI-Nspire tem-se $q = 3$, ou seja $-999 \leq e \leq 999$.

A vantagem do ponto flutuante é que permite representar com o mesmo número de algarismos números de ordem de grandeza muito diferentes. Para representar os números 0.12345 e 1234.5 num sistema sem ponto flutuante (isto é, com “ponto fixo”) são necessários nove dígitos, quatro para a parte inteira (tantos quantos os dígitos da parte inteira do número maior) e cinco para a parte decimal (do número mais próximo de zero). Esta representação será inviável numa máquina que apresente uma limitação em número de dígitos representáveis inferior a nove, mas não o será se a representação admitir que o ponto (ou a vírgula) mude de lugar. Ora, como $0.12345 = 1.2345 \times 10^{-1}$ e $1234.5 = 1.2345 \times 10^3$, uma representação que utilize seis dígitos será suficiente para representar os números de forma exata (neste caso), cinco para a mantissa e um para o expoente.

Apesar da simplicidade deste exemplo, ele revela a razão pela qual as máquinas têm um sistema de ponto flutuante, com ganhos claros ao nível do custo de armazenamento e de processamento. Estão assim preparadas para representar números “muito grandes” e números “muito pequenos” (isto é, próximos de zero), sendo que a mantissa se refere aos algarismos significativos desses números (algarismos da esquerda para a direita e a partir do primeiro diferente de zero) e a grandeza é refletida no expoente.

Descrevemos de seguida a chamada norma IEEE-754 para a representação de números, definida pelo Institute of Electrical and Electronics Engineers em 1985 e revista em 2008 [14]. A norma contempla as representações binárias e decimais. No sistema binário existem o formato simples e o formato duplo com 32 bits e 64 bits, respetivamente. Para além do bit para o sinal do número (0 para positivo e 1 para negativo), há 23 bits para a mantissa e 8 bits para o expoente, no formato simples, e 52 bits para a mantissa e 11 bits para o expoente, no formato duplo. Apesar dos 23 bits para a mantissa, ou 52 no formato duplo, considera-se $p = 24$, ou $p = 53$ no formato duplo, pois a máquina está construída de forma a assumir que o primeiro bit da mantissa é sempre igual a 1, que não ocupa espaço, designando-se por isso de bit implícito.

Por exemplo, o número 19.25, que em binário é 1.001101×2^4 , tem, no formato simples da norma IEEE-754, a representação

$$\begin{array}{ccccccc} 0 & 1000011 & 00110100000000000000000 & & & & \\ | \text{sinal} | & | \text{expoente} | & | \text{mantissa} | & | & | & | & | \end{array}$$

em que não aparece, como se vê, o primeiro bit da mantissa (implícito). A representação dos expoentes é “enviesada” (em inglês, “BIAS exponent”) como a seguir se detalha. Com oito bits, é possível representar todos os números inteiros desde 0 a $255 = 2^8 - 1 = (11111111)_2$ para o expoente, 256 configurações ao todo. Porém, as representações $(00000000)_2$ e $(11111111)_2$ são reservadas para situações excecionais, de underflow e overflow, respetivamente, o que deixa em aberto a utilização de 254 configurações para representar números positivos e também negativos (que são necessários para a representação global de números inferiores à unidade). O sistema não usa bit de sinal e expoentes negativos são obtidos por “subtração” (o enviesamento de que falámos antes) de 127 aos expoentes de 1 até 254. Portanto, desta maneira, é possível representar expoentes na gama de -126 até 127. Isto explica que, no exemplo anterior, o expoente 4 tenha a representação 1000011 (que é de facto a representação não enviesada do número $4+127$).

De um modo geral, qualquer número x normalizado em ponto flutuante com precisão p , pode ser escrito como

$$x = \pm(1.a_1a_2 \dots a_{p-1})_2 \times 2^e.$$

O menor número de \mathcal{F} que é maior do que 1 é $(1.00 \dots 01)_2 = 1 + 2^{-(p-1)}$, sendo $(0.00 \dots 01)_2 = 2^{-(p-1)}$ o que se designa por **epsilon da máquina** (ε). Sendo e o expoente

do número x , $(0.00 \dots 01)_2 \times 2^e$ designa-se por **ulp**(x), abreviatura de “unit in the last place”, isto é, $ulp(x) = \varepsilon \times 2^e$. No exemplo anterior, tem-se $\varepsilon = 2^{-(24-1)} = 2^{-23}$ e $ulp(x) = 2^{-23} \times 2^4 = 2^{-19}$.

Como se disse, numa máquina de calcular gráfica, como as que foram utilizadas neste trabalho, está incorporado um sistema numérico de ponto flutuante $\mathcal{F}(10, p, q)$. Como refere em [15], o expoente é ajustado para que o primeiro dígito da mantissa (à esquerda do ponto decimal) seja diferente de zero e, ao contrário do que sucede com a representação binária, não é possível assumir a existência de um algarismo implícito.

A mantissa de um número representado de forma exata no sistema $\mathcal{F}(10, p, q)$ pode ser escrita como $\pm d_0.d_1d_2 \dots d_{p-1}$, onde, $0 < d_0 < 10$, $0 \leq d_i < 10$ e $i = 0, 1, \dots, p - 1$, sendo p a precisão da máquina.

Assim, a Casio fx-CG20 e a TI-84 Plus C SE têm uma precisão de 13 e o épsilon é 10^{-12} e a precisão da TI-Nspire CX é 14 e o épsilon é 10^{-13} .

Na tabela 2.2 pode observar-se o resultado de um teste para obter o épsilon de cada uma das máquinas utilizadas neste trabalho. Para tal, recorde-se, é necessário subtrair 1 ao menor número de \mathcal{F} que é maior do que 1.

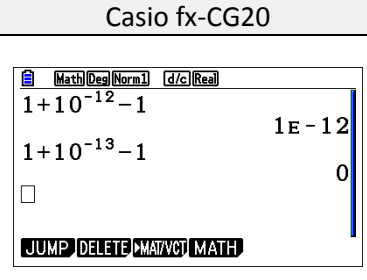
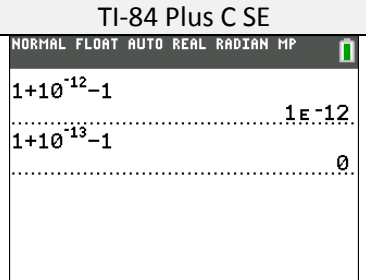
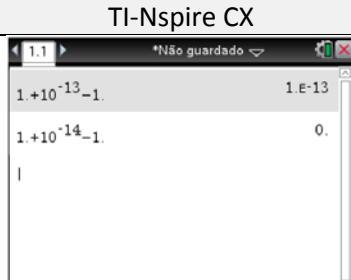
Casio fx-CG20	TI-84 Plus C SE	TI-Nspire CX
		

Tabela 2.2

Quando ao tentar representar um número cujo valor absoluto leva à ultrapassagem do número de dígitos permitido para o expoente, diz-se que é uma situação de **overflow**. Neste caso, é habitual aparecer uma mensagem de erro na calculadora.

Caso contrário, se se tenta representar um número em que o expoente teria de ser inferior ao menor número permitido para o expoente, diz-se que é uma situação de **underflow** e o número representado na máquina é zero.

Vejamos nas figuras 2.15, 2.16 e 2.17, ilustração destas questões relativamente às três máquinas utilizadas neste trabalho.

Casio fx-CG20

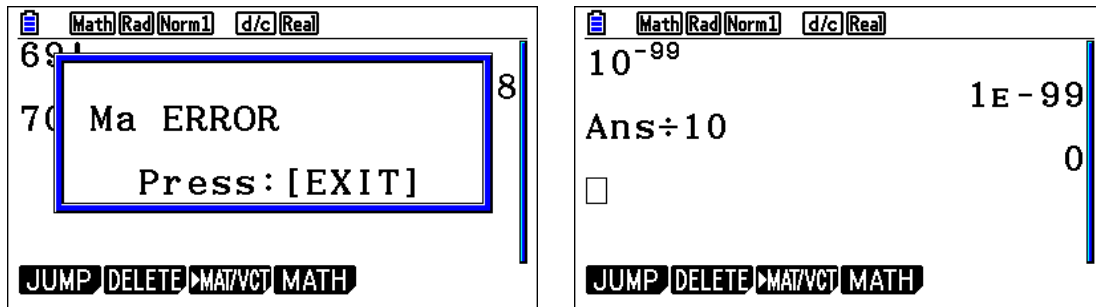


Fig. 2.15

TI-84 Plus C SE

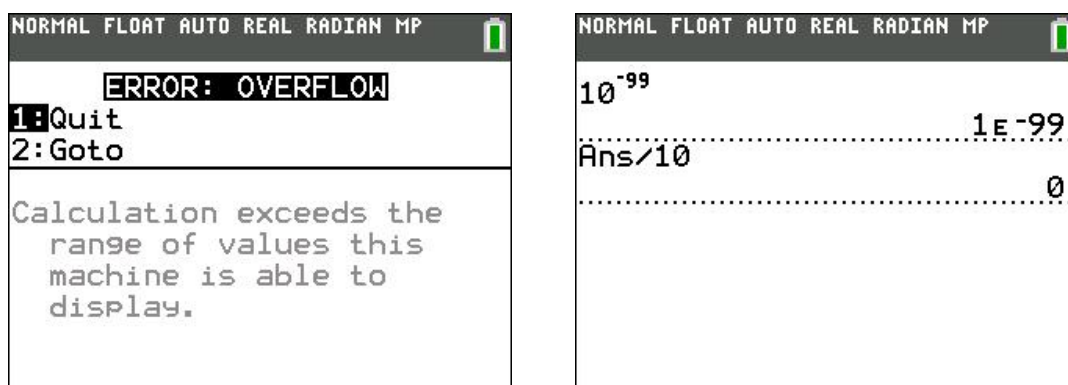


Fig. 2.16

Nota: As mensagens Ma ERROR na Casio fx-CG20 e ERROR: OVERFLOW na TI-84 Plus C SE ocorreram ao tentar calcular $70!$, tendo sido possível obter $69!$ Com expoente +98

TI-Nspire CX

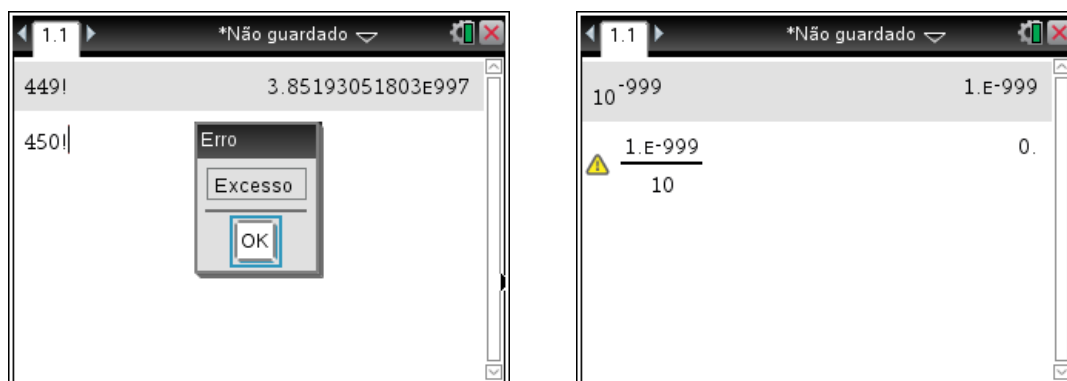


Fig. 2.17

Nota: Nesta calculadora, a situação de underflow surge associada a um símbolo de “perigo”. O perigo é de este falso zero poder afetar e forma catastrófica resultados subsequentes (por exemplo se aparecer como divisor na operação seguinte).

2.3. MODOS DE ARREDONDAMENTO

Como se disse na secção 2.1, dado um número real x (que não causa na máquina problemas de “underflow” ou “Overflow”) será $fl(x) = x$ quando $x \in \mathcal{F}$ e será $fl(x) = x_-$ ou $fl(x) = x_+$ quando $x \notin \mathcal{F}$. A escolha entre estes dois números consecutivos de \mathcal{F} depende do modo de arredondamento e da posição de x no intervalo $[x_-, x_+]$.

Com é referido em [16], um modo de arredondamento possível é o corte (“chopping”), que consiste simplesmente em eliminar os dígitos que estão para além do último admissível. Este modo de arredondamento, também conhecido por arredondamento em direção ao 0 (“rounding toward 0”), é por vezes referido como truncatura. Não lhe chamaremos assim porque, em rigor, um erro de truncatura (“truncation error”) é aquele que se comete quando, por exemplo, se substitui uma soma infinita pela soma de um certo número finito das primeiras parcelas, como se exemplificará na próxima secção.

Por exemplo, ao utilizar este modo de arredondamento para o número

$$\sqrt{3} = 1.73205080756 \dots$$

- num sistema $\mathcal{F}(10,3,2)$, obtém-se 1.73,
- num sistema $\mathcal{F}(10,5,2)$, obtém-se 1.7320, visualizando-se normalmente 1.732 e
- num sistema $\mathcal{F}(10,11,2)$, obtém-se 1.7320508075.

Um modo alternativo usual é o que arredonda para “o mais próximo” e que no caso de ser $x - x_- = x_+ - x$, acrescenta uma unidade na última posição. Para $\sqrt{3}$, este método origina:

- 1.73 num sistema $\mathcal{F}(10,3,2)$,
- 1.7321 num sistema $\mathcal{F}(10,5,2)$ e
- 1.7320508076 num sistema $\mathcal{F}(10,11,2)$.

Além destes dois modos de arredondamento, a norma IEEE-754 prevê ainda os modos de arredondamento, que pelo facto de serem menos usuais apenas se referem a meramente título informativo:

- por defeito, ou no sentido de $-\infty$

O resultado é sempre arredondado para o número representável imediatamente abaixo, isto é, $fl(x) = x_-$.

- por excesso, ou no sentido de $+\infty \rightarrow fl(x) = x_+$

O resultado é sempre arredondado para o número representável imediatamente acima, isto é, $fl(x) = x_+$.

2.4. ERROS NUMÉRICOS

Raras são as vezes em que na vida real é possível calcular soluções exatas para os problemas. Tipicamente, determinamos uma aproximação \bar{x} para a solução exata x .

A qualidade da aproximação \bar{x} pode ser medida pelo **erro absoluto**

$$|x - \bar{x}|$$

ou, como é preferível em muitas situações reais, sendo $x \neq 0$, pelo **erro relativo**

$$\left| \frac{x - \bar{x}}{x} \right|$$

o qual se expressa muitas vezes em termos de uma percentagem. Observe-se que, sendo conhecida a aproximação \bar{x} mas não o valor exato x , não é possível determinar os valores dos erros definidos antes. Na prática, o que se faz é encontrar valores o mais pequenos possível que majoram os erros.

Exemplo 2.1

Não é pouco frequente a utilização do valor aproximado 3.14 para π . Já no passado se chegou a aproximações deste número transcendente como 3 (1009 a 922 a.C. - I Reis 7:23 da Bíblia Sagrada) ou $\frac{22}{7}$ (gregos - 250 a.C.). Analisemos na tabela 2.3 este tipo de erros para estas aproximações:

$\bar{\pi}$	$ \pi - \bar{\pi} $	$\left \frac{\pi - \bar{\pi}}{\pi} \right $
3	0.14159 ...	0.04507 ...
3.14	0.01592 ...	0.0005069 ...
$\frac{22}{7}$	0.00126 ...	0.0004024 ...

Tabela 2.3

Exemplo 2.2

$x = \ln 2 = 0.66314718 \dots$, com $\bar{x} = 0.6$.

$$\text{erro absoluto} = |\ln 2 - 0.6| = 0.093 \dots \quad \text{erro relativo} = \left| \frac{\ln 2 - 0.6}{\ln 2} \right| = 0.134 \dots$$

Exemplo 2.3

O número de Avogadro (número de moléculas numa mole - $6.02214129 \times 10^{23} \text{mol}^{-1}$) é muitas vezes usado com a aproximação $6.02 \times 10^{23} \text{mol}^{-1}$. Neste caso é gerado um erro absoluto bastante elevado de cerca de $2.14 \times 10^{20} \text{mol}^{-1}$, mas um erro relativo inferior a 0.04% ($3.55 \dots \times 10^{-4}$).

Nos exemplos seguintes ilustramos erros de truncatura.

Exemplo 2.5

A partir do conhecido desenvolvimento

$$\ln 2 = S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \dots$$

podemos obter uma representação decimal de $\ln 2$ tão precisa quanto se desejar, bastando para tal adicionar um número suficientemente grande (obviamente finito) dos primeiros termos. A soma dos termos desprezados é, de acordo com a definição dada antes, um erro de truncatura.

Sendo a soma dos primeiros seis termos igual a $\frac{37}{60} = 0.61(6)$, o erro de truncatura é $\frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} + \dots$, ou seja cerca de 0.07648.

Exemplo 2.6

As calculadoras gráficas (que não possuam capacidades de cálculo algébrico) usam fórmulas numéricas para o cálculo aproximado de derivadas (isto é, fórmulas de derivação numérica).

Uma dessas fórmulas, conhecida pela fórmula das diferenças centrais (1) será a que utiliza a calculadora gráfica TI-84 Plus C SE para obter a derivada de uma função f num ponto.

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} \quad (1)$$

Uma vez que a derivada $f'(x)$ é o limite quando h tende para zero do quociente anterior, a derivação numérica consiste em substituir o cálculo do limite (isto é, onde h é um infinitésimo) pelo valor obtido no cálculo (finito) do quociente para um valor de h pequeno. Um problema da fórmula (1) é que, usada cegamente, ela dará zero para o valor da derivada da função módulo no ponto $x = 0$, como se pode observar na figura 2.19.

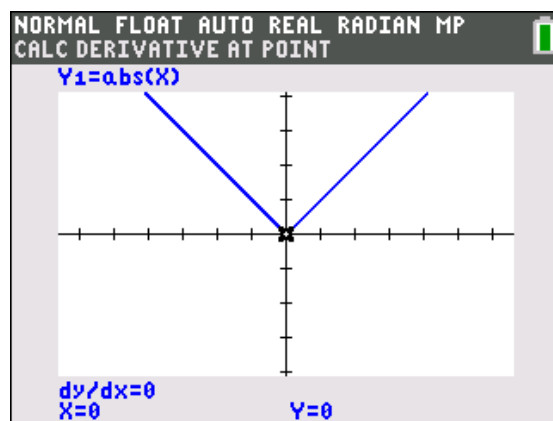


Fig. 2.19

2.4.1. ERROS DE ARREDONDAMENTO

Relativamente aos erros cometidos na representação da máquina, é importante perceber de que forma dependem da precisão p e do próprio número arredondado. Vamos considerar a base decimal, uma vez que é a utilizada pelas máquinas de calcular.

Para um número real na forma $x = \pm m \times 10^e$ e a sua representação $fl(x) = \pm \tilde{m} \times 10^e$, vamos definir, em coerência com o referido na secção anterior, o erro absoluto por $A = |\text{round}(x) - x|$. Assim, tem-se que $A = |\tilde{m} - m| \times 10^e$.

Independentemente do modo de arredondamento utilizado, tem-se

$$A < \varepsilon \times 10^e = \text{ulp}(x) \quad (2)$$

e no caso do arredondamento para o mais próximo é

$$A < \frac{\text{ulp}(x)}{2} \quad (3)$$

Designando por R o erro relativo $\frac{|\tilde{x}-x|}{|x|}$, tem-se que $R = \frac{|\tilde{m}-m|}{|m|}$ e uma vez que $|m| \geq 1$ resulta

$$R < \varepsilon \quad (4)$$

no geral, ou

$$R < \frac{1}{2} \varepsilon \quad (5)$$

no caso de arredondamento para o mais próximo.

O que se deve destacar destas expressões é que o majorante dado em (2) para o erro absoluto de arredondamento cresce com o próprio número (isto é, com o expoente e do número) mas o majorante dado em erro relativo nunca é maior do que o ε da máquina, independentemente da grandeza do número e esta é uma das consequências mais importantes deste modo de representação.

Na tabela 2.4 ilustra-se, em certa medida, o que se acabou de dizer. Na representação de três números num sistema $\mathcal{F}(10,4,2)$, pode observar-se que o mesmo erro relativo pode corresponder a erros absolutos totalmente distintos, fazendo do erro relativo uma medida mais significativa do que o erro absoluto porque toma em consideração a grandeza do valor exato, aliás como já se viu na secção anterior em que o contexto era mais alargado do que a aproximação por arredondamento num sistema de ponto flutuante.

x	\tilde{x}	A	R
3.1415×10^{-3}	3.142×10^{-3}	5×10^{-7}	$1.59 \dots \times 10^{-4}$
3.1415×10^0	3.142×10^0	5×10^{-4}	$1.59 \dots \times 10^{-4}$
3.1415×10^5	3.142×10^5	5×10^1	$1.59 \dots \times 10^{-4}$

Tabela 2.4

O número de algarismos (ou dígitos) significativos dum valor aproximado dá boa ideia da qualidade dessa aproximação.

No geral, diz-se que $fl(x)$ é um valor aproximado de x , arredondado para k casas decimais corretas, se e só se

$$A = |fl(x) - x| \leq 0.5 \times 10^{s+1-k} \quad (6)$$

sendo s tal que $10^s \leq |x| < 10^{s+1}$.

De (3), vem que, num sistema decimal de ponto flutuante $R = \left| \frac{fl(x)-x}{x} \right| \leq \frac{0.5 \times 10^{s+1-k}}{10^s}$, donde

$$R = \left| \frac{fl(x) - x}{x} \right| \leq 0.5 \times 10^{1-k} \quad (7)$$

Assim, há também uma relação entre o número de algarismos significativos dum valor aproximado dum número e o erro relativo desse valor [17]. Vejamos, a este propósito, uma concretização das fórmulas (6) e (7) com a aproximação 2.742×10^{-3} de 2.7423×10^{-3} .

Ora, tem-se que $s = -3$, $A = 3 \times 10^{-7}$ e $R = 1.09 \dots \times 10^{-4} \leq 0.5 \times 10^{-3}$. De (6) vem que $-3 + 1 - k = -6$ e de (7) vem que $1 - k = -3$, o que leva a $k = 4$ em ambas as situações, pelo que a conclusão confirma o que sabemos, ou seja, que o número arredondado coincide com o valor exato nos primeiros quatro algarismos significativos.

2.4.2. PROPAGAÇÃO DOS ERROS NAS OPERAÇÕES ARITMÉTICAS

Nesta secção estudamos abreviadamente a propagação dos erros nas quatro operações aritméticas fundamentais. Esta situação suscita a necessidade de dedicar um pouco mais de atenção às operações aritméticas no sistema de ponto flutuante, mais conhecidas por operações aritméticas de ponto flutuante [15, p. 77]. Estas operações substituem as operações aritméticas usuais porque não se podem fazer exatamente num computador, e como tal há também erros associados às operações. Note-se que num sistema destes o resultado de qualquer das operações $+$, $-$, \times e \div entre dois números de \mathcal{F} pode não pertencer a \mathcal{F} . Neste caso, a norma IEEE-754 requer que o resultado produzido seja o valor corretamente arredondado do resultado exato, ou seja, arredondado de modo a pertencer ao formato utilizado, através do modo de arredondamento

definido. Por exemplo, num sistema $\mathcal{F}(10,4,2)$ se $x_1 = 1.234 \times 10^1$ e $x_2 = 6.437 \times 10^{-1}$, $x_1 + x_2 = 1.8777 \times 10^1$, que não pertence a \mathcal{F} e que depois de arredondado fica 1.878×10^1 . Outro aspeto interessante é o facto de que as propriedades usuais das operações não são também todas válidas num sistema de ponto flutuante. Por exemplo, ao calcular $0.5 + 0.24 + 0.012$ num sistema $\mathcal{F}(10,2,1)$, que na aritmética exata é 0.536 , temos:

- adicionando da esquerda para a direita, vem:

$$(0.50 \times 10^0 + 0.24 \times 10^{-1}) + 0.12 \times 10^{-1} =$$

$$(0.50 \times 10^0 + 0.02 \times 10^0) + 0.12 \times 10^{-1} = 0.52 \times 10^0 + 0.01 \times 10^0 = 0.53 \times 10^0$$

- adicionando da direita para a esquerda, vem:

$$0.50 \times 10^0 + (0.24 \times 10^{-1} + 0.12 \times 10^{-1}) =$$

$$0.50 \times 10^0 + 0.36 \times 10^{-1} = 0.52 \times 10^0 + 0.04 \times 10^0 = 0.54 \times 10^0$$

Comprova-se assim que a propriedade associativa da adição não é válida com a aritmética num sistema de ponto flutuante.

As operações efetuadas nos exemplos são a concretização das expressões gerais que se podem observar na tabela 2.5 ao operar dois números $x_1 = m_1 \times b^{e_1}$ e $x_2 = m_2 \times b^{e_2}$, com $b = 10$,

Adição e subtração $\rightarrow x_1 \pm x_2$ $= \begin{cases} (m_1 \pm m_2 \times b^{-(e_1-e_2)}) \times b^{e_1} & \text{se } e_1 > e_2 \\ (m_1 \times b^{-(e_2-e_1)} \pm m_2) \times b^{e_2} & \text{se } e_1 \leq e_2 \end{cases}$	
Multiplicação $\rightarrow x_1 \times x_2$ $= (m_1 \times m_2) \times b^{e_1+e_2}$	Divisão $\rightarrow \frac{x_1}{x_2}$ $= \left(\frac{m_1}{m_2}\right) \times b^{e_1-e_2}$

Tabela 2.5

e os procedimentos resultam do facto destas operações se desenvolverem em computador de acordo com os seguintes trâmites [19]:

1. Decomposição dos operandos, isto é, separação destes números nas respetivas mantissas e expoentes;
2. Alinhamento das mantissas e dos expoentes para a soma e para a subtração;
3. Operação com as mantissas e/ou com os expoentes;
4. Normalização da mantissa, isto é, translação à esquerda da mantissa com decréscimo do expoente em uma unidade por cada zero à esquerda da mantissa;
5. Arredondamento da mantissa.
6. Recomposição do resultado, isto é, reunião da mantissa e do expoente para formar o resultado no sistema de ponto flutuante.

Pode dizer-se que se x e y são números no sistema de ponto flutuante:

$$x \oplus y = fl(fl(x) + fl(y))$$

$$x \ominus y = fl(fl(x) - fl(y))$$

$$x \otimes y = fl(fl(x) \times fl(y))$$

$$x \oslash y = fl(fl(x) \div fl(y))$$

O facto de se conhecer com algum detalhe a aritmética do ponto flutuante permite estar melhor preparado para enfrentar algumas armadilhas que podem surgir nos cálculos de computador [16, p. 16]. Assim, veremos como estas operações podem levar à propagação de erros e analisaremos em concreto um efeito importante decorrente desta propagação de erros ao longo das operações, o cancelamento subtrativo. A propagação de erros não se deve aos expoentes, por se tratarem de números inteiros (exceto situações de underflow e overflow, que não serão aqui consideradas), mas a ocorrerem será devido à mantissa.

Tendo em consideração o que foi referido na secção anterior, ao considerar o arredondamento para o mais próximo e designando por δ_x o erro absoluto de arredondamento de $fl(x)$, pode escrever-se $fl(x) = x + \delta_x$, sabendo-se que $\delta_x \leq \frac{1}{2} ulp(x)$ e ainda $fl(x) = x(1 + \delta)$, em que $|\delta| \leq \frac{1}{2} \varepsilon$.

Consideremos x e y dois números reais e \tilde{x} e \tilde{y} as respetivas representações num sistema de ponto flutuante. Vejamos, por exemplo, como se propagam os erros na multiplicação de x por y . Seja $z = x \times y$ o produto exato e $fl(z) = x \otimes y$ o resultado obtido por arredondamento para o mais próximo no sistema de ponto flutuante.

Então,

$$\begin{aligned} fl(z) &= [x(1 + \delta_1)y(1 + \delta_2)](1 + \delta_3) = xy(1 + \delta_1)(1 + \delta_2)(1 + \delta_3) = \\ &= xy(1 + \delta_1 + \delta_2 + \delta_3 + \delta_1\delta_2 + \delta_1\delta_3 + \delta_2\delta_3 + \delta_1\delta_2\delta_3) \\ &\approx xy + xy(\delta_1 + \delta_2 + \delta_3), \text{ uma vez } \delta_i \text{ são já muito pequenos e os produtos} \end{aligned}$$

envolvendo o multiplicação de mais do que um destes valores podem ser ignorados.

Assim, nota-se que o erro relativo é praticamente o mesmo do que se obteria com os operandos perturbados. O erro absoluto ao multiplicar dois números não é superior a uma vez e meia o épsilon da máquina, pois

$$A = |xy(\delta_1 + \delta_2 + \delta_3)| \leq |xy|(|\delta_1| + |\delta_2| + |\delta_3|) \leq |x||y| \times 3 \frac{\varepsilon}{2}$$

Na realidade, na multiplicação os erros relativos não se propagam rapidamente, podendo verificar-se que também na divisão isso acontece [20, p. 24]. Vejamos agora o que sucede com a adição. Seja $s = x + y$ a soma exata e $fl(s) = x \oplus y$ a soma no sistema de ponto flutuante após arredondamento para o mais próximo.

Assim,

$$\begin{aligned} fl(s) &= [x(1 + \delta_1) + y(1 + \delta_2)](1 + \delta_3) = \\ &= x(1 + \delta_1 + \delta_3 + \delta_1\delta_3) + y(1 + \delta_2 + \delta_3 + \delta_2\delta_3) \\ &\approx x(1 + \delta_1 + \delta_3) + y(1 + \delta_2 + \delta_3) \\ &= x + y + x(\delta_1 + \delta_3) + y(\delta_2 + \delta_3) \end{aligned}$$

O erro absoluto neste cálculo é

$$A = |x(\delta_1 + \delta_3) + y(\delta_2 + \delta_3)| \leq |x|(|\delta_1| + |\delta_3|) + |y|(|\delta_2| + |\delta_3|) \leq (|x| + |y|)\varepsilon$$

Este erro é da ordem do produto de ε por qualquer valor $|x|$ ou $|y|$.

O erro relativo é:

$$\frac{|x|\varepsilon + |y|\varepsilon}{|x + y|} \quad (8)$$

Se $|x + y|$ é próximo de zero, ou seja, x ou y são muito próximos do simétrico do outro (se considerássemos a operação de subtração tal ocorreria com x e y muito próximos), havendo por isso perda de algarismos significativos, o erro relativo pode ser muito elevado comparativamente ao de cada operando, resultado daí um erro que se chama de **cancelamento subtrativo**, com efeitos muito relevantes no resultado da operação.

Observe-se um exemplo da propagação do erro na subtração com consequência no cancelamento subtrativo. Suponhamos que depois de alguns cálculos se obtêm dois números corretos com 8 algarismos significativos e agora é necessário subtraí-los. Por exemplo:

$$\begin{array}{r} 21.234568 \\ - 21.234567 \\ \hline 00.000001 \end{array}$$

Neste caso, as parcelas são dadas com 8 algarismos significativos de precisão e a diferença tem uma precisão de apenas 1 algarismo significativo. Este exemplo é relativamente simples, mas há outras situações de perda de significado nas subtrações de números muito próximos que são mais subtis e que em muitos casos é possível evitá-los com reformulação do cálculo.

De seguida serão analisados alguns exemplos em que o cancelamento subtrativo leva a erros eventualmente indesejáveis e a forma como se podem evitar ou reduzir com algumas reformulações do cálculo.

Exemplo 2.7 [16, p. 16]

Consideremos o cálculo de $\sqrt{9.01} - 3$ numa máquina em que os números que opera têm três algarismos significativos. Sabe-se que o valor exato de $\sqrt{9.01}$ é, com 4 algarismos significativos, 3.002, pelo que a máquina referida utiliza 3.00.

Assim, $\sqrt{9.01} - 3 = 3.00 - 3 = 0.00$, que não tem dígitos significativos e cuja diferença real não é claramente a que obtivemos. O problema é que estamos a subtrair números muito próximos, pelo que uma transformação suscitada por uma igualdade algébrica pode levar a evitar-se esta subtração e ao mesmo tempo o erro será reduzido.

Note-se que $a - b = \frac{a^2 - b^2}{a + b}$. Aplicando esta propriedade à nossa subtração, vem que:

$$\sqrt{9.01} - 3 = \frac{0.01}{\sqrt{9.01} + 3} = \frac{0.01}{6.00} = 0.001(6) = 1.67 \times 10^{-3}, \text{ muito mais próximo do valor exato.}$$

Na figura 2.20 podemos observar um efeito deste tipo de problema numa calculadora gráfica (TI-84 Plus C SE) e como se obtém, analogamente, a diferença com mais rigor.

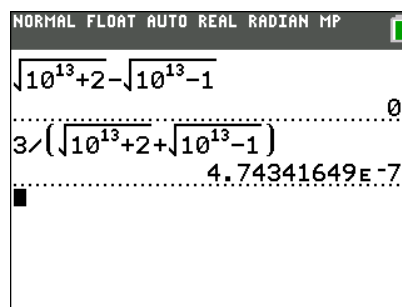


Fig. 2.20

Exemplo 2.8

Vamos resolver a equação $x^2 + 79.05x + 1 = 0$ numa “máquina” com um sistema que utiliza 4 dígitos significativos e também numa “máquina” em que são utilizados 5 dígitos significativos.

Note-se que uma equação do tipo $ax^2 + bx + c = 0$, $a \neq 0$ se pode resolver com recurso à fórmula resolvente, em que as raízes se podem obter fazendo:

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \text{ e } x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad (9)$$

Se procurarmos as raízes exatas e as apresentarmos com 6 dígitos significativo, vem que:

$$x_1 = -79.0373 \text{ e } x_2 = -0.0126522.$$

Utilizando as fórmulas em (9), na máquina de 4 dígitos, vem que:

$$fl(x_1) = \frac{-79.05 - \sqrt{6249 - 4}}{2} = \frac{-79.05 - 79.03}{2} = -79.04$$

$$fl(x_2) = \frac{-79.05 + \sqrt{6249 - 4}}{2} = \frac{-79.05 + 79.03}{2} = -0.01$$

Vejamos qual foi o erro relativo cometido em cada um dos casos, respetivamente:

$$\left| \frac{-79.04 - (-79.04)}{-79.04} \right| = 0$$

e

$$\left| \frac{-0.01 - (-0.01265)}{-0.01265} \right| \approx 0.21$$

O erro relativo na determinação de $fl(x_2)$ é relevante.

Mais uma vez este erro ocorre devido ao cancelamento subtrativo e pode também ser reduzido ao efetuar uma reformulação no modo de obter as soluções.

Uma equação geral do 2º grau, $ax^2 + bx + c = 0$, $a \neq 0$, pode reescrever-se na forma:

$$x^2 - \left(-\frac{b}{a}\right)x + \frac{c}{a} = 0, \text{ em que } -\frac{b}{a} = x_1 + x_2 \text{ e } \frac{c}{a} = x_1 \times x_2,$$

(x_1 e x_2 são soluções da equação)

Ora, temos que $a = 1$, $c = 1$ e sabe-se que -79.04 é a raiz obtida sem ter havido cancelamento subtrativo.

Da igualdade $\frac{c}{a} = x_1 \times x_2$, vem que: $fl(x_2) = \frac{1}{-79.04} = -0.01265$, agora com erro relativo

nulo, pois $\left| \frac{-0.01265 - (-0.01265)}{-0.01265} \right| = 0$.

Se tivéssemos utilizado uma máquina com 5 dígitos significativos, teríamos:

$$fl(x_1) = \frac{-79.05 - \sqrt{6249 - 4}}{2} = \frac{-79.05 - 79.025}{2} = -79.038$$

$$fl(x_2) = \frac{-79.05 + \sqrt{6249 - 4}}{2} = \frac{-79.05 + 79.025}{2} = -0.0125$$

Com erros relativos, respetivamente,

$$\left| \frac{-79.038 - (-79.037)}{-79.037} \right| \approx 1.3 \times 10^{-5}$$

e

$$\left| \frac{-0.0125 - (-0.012652)}{-0.012652} \right| \approx 0.012$$

temos novamente um erro relativo consideravelmente superior para a solução de menor valor absoluto, quase mil vezes superior ao da solução de maior valor absoluto.

$$fl(x_2) = \frac{1}{-79.038} = -0.012652 \text{ tem agora erro relativo nulo, pois } \left| \frac{-0.012652 - (-0.012652)}{-0.012652} \right| = 0.$$

Vejamos este efeito com a utilização de uma calculadora gráfica, agora a Casio fx-CG20:

Exemplo 2.9

Para resolver a equação $x^2 + 9^{12}x - 3 = 0$ [16, p. 17], a Casio utiliza funcionalidades internas que contemplam técnicas para evitar erros como o do cancelamento subtrativo, o que é comum a outras calculadoras, que têm essas funcionalidade em paralelo com a página de cálculos aritméticos. Os algoritmos destas funcionalidades não são dados a conhecer pelo fabricante mas pode observar-se o seu efeito nas figuras 2.21 a 2.23.

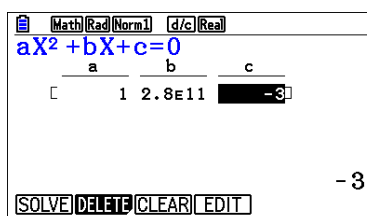


Fig. 2.21

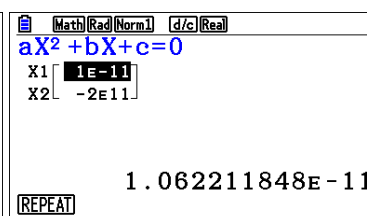


Fig. 2.22

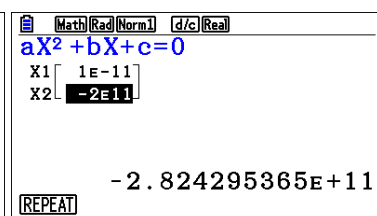


Fig. 2.23

Em [16, p. 17] refere que utilizando o MatLab com precisão de 16 dígitos dá 0 para a solução mais próxima de 0, tal como o que sucede com a Casio fx-CG20, se agora for utilizada uma página de cálculos, como se pode observar na figura 2.24.

The screenshot shows a Casio calculator interface with the following text and symbols:

- Top bar: **Math** **Rad** **Norm1** **d/c** **Real**
- Equation 1: $(-9^{12} - \sqrt{9^{24} + 12}) \div 2$
- Equation 2: $(-9^{12} + \sqrt{9^{24} + 12}) \div 2$
- Result: $-2.824295365E+11$
- Final result: 0
- Bottom bar: **JUMP** **DELETE** **MAT/VCT** **MATH**

Fig. 2.24

Efetuada uma mudança da forma de cálculo dessa solução, a mesma que foi realizada no exemplo 2.8, podemos obter um valor muito ajustado para essa solução, tal como o que se obteve na referida obra e que se pode observar na figura 2.25.

The screenshot shows a Casio calculator interface with the following text and symbols:

- Top bar: **Math** **Rad** **Norm1** **d/c** **Real**
- Equation 1: $(-9^{12} - \sqrt{9^{24} + 12}) \div 2$
- Equation 2: $(-9^{12} + \sqrt{9^{24} + 12}) \div 2$
- Result: $-2.824295365E+11$
- Equation 3: $-3 \div (-2.824295365E+11)$
- Result: $1.062211848E-11$
- Final result: 0
- Bottom bar: **JUMP** **DELETE** **MAT/VCT** **MATH**

Fig. 2.25

Para Maria Raquel Valença, na sua obra [21], o objetivo da Análise Numérica é o estudo de métodos que permitem resolver problemas matemáticos usando operações de aritmética. Refere ainda que por essa razão os métodos numéricos são particularmente adequados para resolver problemas de cálculo em computadores e que a compreensão e a análise dos efeitos da representação finita de números e processos pelos computadores passou a fazer parte dos objetivos da análise numérica por causa da problemática da representação de números reais num conjunto finito.

Na realidade, neste trabalho, o foco principal é a análise de alguns métodos numéricos para resolver equações, as quais podem estar ou não relacionadas com a resolução de problemas mais ligados a situações de modelação, mas foi importante ter analisado a aritmética de cálculo em computadores para no próximo capítulo se passar à análise de alguns desses métodos de resolução de equações.

3. ALGUNS MÉTODOS NUMÉRICOS DE RESOLUÇÃO DE EQUAÇÕES

3.1. INTRODUÇÃO

Encontrar as soluções de uma equação é um dos problemas mais importantes e mais antigos da matemática e de outras ciências. Como se pode concluir do que é referido em [22, p. 48], e em [16, p. 39], problemas deste género apareceram há pelo menos 3700 anos. Numa tábua de argila com escrita cuneiforme, que consta da coleção de documentos babilónios da Universidade de Yale, nos Estados Unidos, datada de cerca de 1700 A.C., YBC7289 (figura 3.1), pode observar-se, em escrita sexagesimal, uma aproximação de $\sqrt{2}$ como a medida da diagonal de um quadrado de lado unitário, a qual foi obtida com 5 casas decimais depois de serem efetuados três passos num procedimento repetitivo.

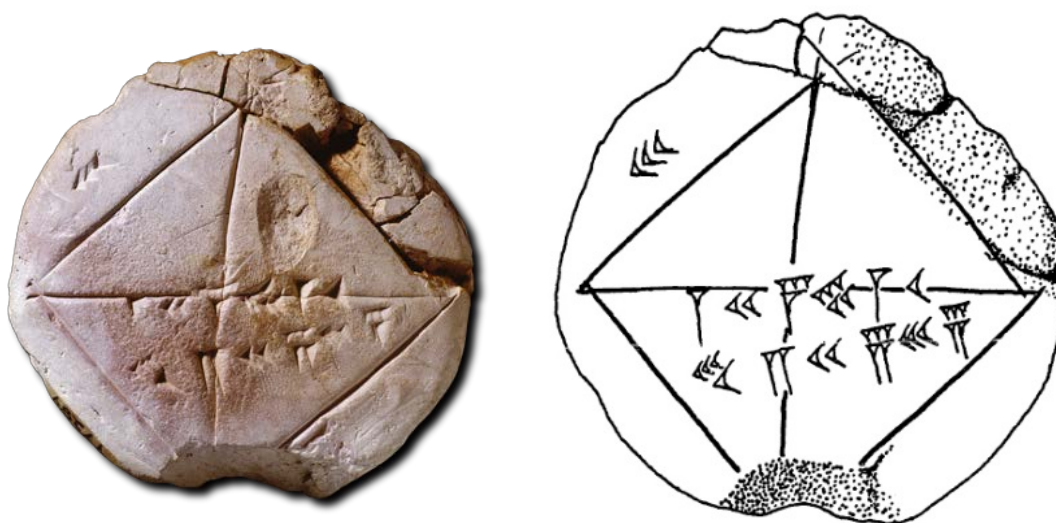


Fig. 3.1

(Imagens obtidas em [23] e em [24])

A qualquer momento pode ser necessário resolver uma equação, mas na realidade não há métodos diretos que nos permitam encontrar as soluções de todas as equações num número finito de passos. Podem resolver-se certas equações algébricas, embora não haja fórmulas para equações polinomiais de grau superior a 5 [25], e algumas exponenciais, logarítmicas ou

trigonométricas, mas não todas. Não há métodos, por exemplo, para obter todas as soluções exatas da equação $x^2 = 2^x$.

Que fazer então? Não temos alternativa que não seja a de procurar aproximações das soluções, tão boas quanto necessário e/ou quanto possível. Serão vistos de seguida três métodos, cuja seleção a partir de um número elevado de métodos se deve ao facto de serem muito utilizados e suscitarem questões pertinentes para tratamento ao nível do ensino não superior. Trata-se do método da bissecção, do método do ponto fixo e do método de Newton-Raphson.

3.2. MÉTODO DA BISSEÇÃO

3.2.1. INTRODUÇÃO

Em [16, p. 25], Sauer faz uma analogia entre este método iterativo de resolução de equações e a procura de um nome numa lista telefónica, o que também poderia ser comparável à procura de uma palavra num dicionário em papel. Por exemplo, como procurar a palavra “bisseção” num desses dicionários? Em princípio abre-se o dicionário nas primeiras páginas, pois “B” é a 2ª letra do alfabeto, sendo esta uma primeira tentativa de aproximação à palavra que se procura. Poderíamos ter encontrado uma página com palavras começadas pela letra A, antes portanto, da página que interessa. De seguida, folheamos um de uma só vez um conjunto de páginas e chegamos, por exemplo, a uma página com a referência BRO, onde pelo menos algumas palavras começam por este terno de letras, portanto depois da página que interessa. Define-se assim, e de forma implícita, um intervalo inicial onde podemos encontrar a palavra pretendida. De seguida, é natural que se vá reduzindo a amplitude deste intervalo e que a divisão do conjunto de páginas mais ou menos ao meio seja uma estratégia plausível. Se nessa divisão verificarmos que “caímos” antes da palavra, esta página constitui o novo extremo esquerdo do intervalo, caso contrário, a página considerada constitui o novo extremo direito do intervalo, e assim sucessivamente até encontrar a página da palavra, e depois a palavra de entre as que estão por ordem alfabética nessa página.

Acautelando as devidas diferenças, esta comparação diz bem da natureza deste método.

Para procurar uma solução de uma equação do tipo $f(x) = 0$, seguindo este método, é importante começar a tentar localizá-la, definindo-se um primeiro intervalo $[a, b]$ onde ela se poderá encontrar. É uma boa opção verificar o sinal de f em a e em b , pois se forem contrários, é garantido, pelo corolário do teorema de Bolzano, que há pelo menos uma solução da equação no intervalo, desde que f seja uma função contínua nesse intervalo.

3.2.2. FUNCIONAMENTO DO MÉTODO

O funcionamento deste método baseia-se no seguinte teorema.

Teorema 3.1. – Corolário do Teorema de Bolzano

Seja f uma função contínua em $[a, b]$. Se $f(a) \times f(b) < 0$, então a função f tem pelo menos uma raiz em $]a, b[$, ou seja, $\exists r \in]a, b[: f(r) = 0$.

Consideremos, por uma questão de simplicidade, como se faz em [22, p. 48], que há apenas uma solução da equação $f(x) = 0$ num intervalo $]a, b[$, sendo f uma função contínua definida em $[a, b]$ e $f(a) \times f(b) < 0$, pelo que a existência desta solução é garantida pelo teorema 3.1.

A técnica da bisseção consiste em obter, a partir de $[a, b]$, intervalos de amplitude cada vez menor onde se localize a solução. Para colocar o método em funcionamento começa-se por utilizar o valor central p de $[a, b]$, que será extremo de um subintervalo $[a, p]$ ou $[p, b]$ onde se localize a solução, o que é decidido com aplicação do teorema 3.1. Repete-se depois o procedimento sobre esse intervalo e sucessivamente até que o valor central seja a solução ou se chegue a um intervalo que tenha uma amplitude até um erro admissível à partida para a solução.

Como refere em [22, p. 48], considera-se inicialmente como extremos do intervalo $a_1 = a$ e $b_1 = b$ e calcula-se o valor central do intervalo, $p_1 = a_1 + \frac{b_1 - a_1}{2} = \frac{a_1 + b_1}{2}$.

Se $f(p_1) = 0$,

temos que p_1 é solução exata da equação $f(x) = 0$.

Se $f(p_1) \neq 0$,

$f(p_1)$ tem o mesmo sinal de $f(a_1) \rightarrow f(a_1) \times f(p_1) > 0$

ou de $f(b_1) \rightarrow f(a_1) \times f(p_1) < 0$.

Neste caso,

- Se $f(a_1) \times f(p_1) > 0$, faça-se $a_2 = p_1$ e $b_2 = b_1$.
- Se $f(a_1) \times f(p_1) < 0$, faça-se $a_2 = a_1$ e $b_2 = p_1$.

Repete-se de seguida este algoritmo com $[a_2, b_2]$ e assim sucessivamente até que se tenha algum p_i , tal que $f(p_i) = 0$ ou o valor central p_i esteja tão próximo quanto necessário da solução.

A figura 3.2 [22, p. 49], ilustra este método iterativo de determinação da solução ou de aproximação à solução, até à 3ª estimativa.

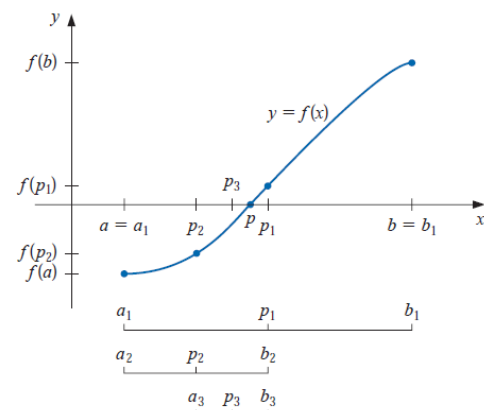


Fig. 3.2

Observe-se a aplicação deste método a um exemplo particular.

Exemplo 3.1

Considere-se a equação $x^2 = 2$, a qual se pode reescrever na forma $x^2 - 2 = 0$ de modo a possibilitar a aplicação do método da bisseção utilizando a função definida por $f(x) = x^2 - 2$.

Uma folha de cálculo com alguma programação nas células de modo a desenvolver o algoritmo que constitui o processo da bisseção facilita um trabalho repetitivo e que pode ser longo.

Na tabela 3.1 pode observar-se o desenvolvimento do método com a apresentação de 10 iterações, desde o intervalo $[1,2]$. Este é um intervalo aceitável para iniciar o processo, uma vez que $f(1) \times f(2) = (-1) \times 2 < 0$, garantindo o teorema 3.1 a existência da solução.

i	a_i	b_i	p_i	Sinal de $f(a_i) \times f(p_i)$
1	1,0000000000	2,0000000000	1,5000000000	-
2	1,0000000000	1,5000000000	1,2500000000	+
3	1,2500000000	1,5000000000	1,3750000000	+
4	1,3750000000	1,5000000000	1,4375000000	-
5	1,3750000000	1,4375000000	1,4062500000	+
6	1,4062500000	1,4375000000	1,4218750000	-
7	1,4062500000	1,4218750000	1,4140625000	+
8	1,4140625000	1,4218750000	1,4179687500	-
9	1,4140625000	1,4179687500	1,4160156250	-
10	1,4140625000	1,4160156250	1,4150390625	-

Tabela 3.1

Instrução de preenchimento da célula $a_i, i > 1$:

Se $f(a_{i-1}) \times f(p_{i-1}) < 0$, escreve a_{i-1} , senão escreve p_{i-1}

Instrução de preenchimento da célula $b_i, i > 1$:

Se $f(a_{i-1}) \times f(p_{i-1}) < 0$, escreve p_{i-1} , senão escreve b_{i-1}

Na figura 3.3 pode observar-se graficamente a evolução dos valores centrais p_i . Estes valores parecem estabilizar entre 1.40 e 1.42, que na realidade é a solução da equação $x^2 - 2 = 0$, $\sqrt{2}$. ($\sqrt{2} = 1.414213562 \dots$)

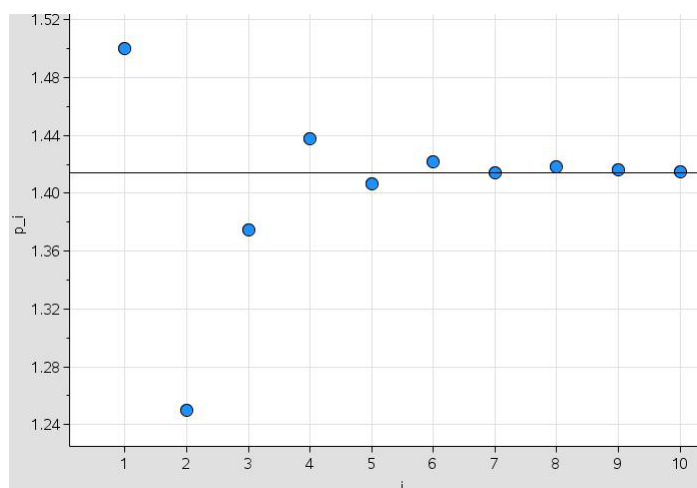


Fig. 3.3

Continuando a efetuar mais algumas iterações podemos reforçar a conjectura da estabilização, o que se pode observar na tabela 3.2 e na figura 3.4..

i	a_i	b_i	p_i	S
11	1,414062500000	1,415039062500	1,414550781250	-
12	1,414062500000	1,414550781250	1,414306640625	-
13	1,414062500000	1,414306640625	1,4141845703125	+
14	1,4141845703125	1,414306640625	1,4142456054688	-
15	1,4141845703125	1,4142456054688	1,4142150878906	-
16	1,4141845703125	1,4142150878906	1,4141998291016	+
17	1,4141998291016	1,4142150878906	1,4142074584961	+
18	1,4142074584961	1,4142150878906	1,4142112731934	+
19	1,4142112731934	1,4142150878906	1,4142131805420	+
20	1,4142131805420	1,4142150878906	1,4142141342163	-
21	1,4142131805420	1,4142141342163	1,4142136573792	-
22	1,4142131805420	1,4142136573792	1,4142134189606	+

Tabela 3.2

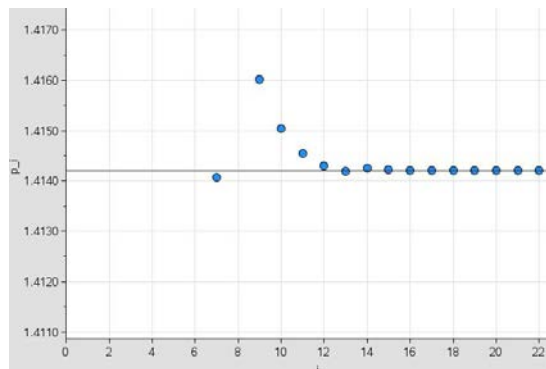


Fig. 3.4

Importa analisar como evolui a diferença entre os valores centrais, que se podem considerar valores aproximados da solução, e o valor exato, ou seja, a medida do erro absoluto, à medida que se fazem iterações. Com esta análise poderemos concluir sobre o número de iterações necessárias para que o método devolva uma solução com uma aproximação requerida.

Todo este processo deixa de fazer sentido a partir do momento em que se sabe o valor exato da solução. Não o sabendo, não é possível calcular, em cada iteração, o erro absoluto ou o erro relativo. Mas como se conhece a amplitude de cada intervalo onde a solução se encontra, é possível majorar o valor do erro. Assim, podem obter-se valores aproximados da solução tão precisos quanto se pretenda, dentro das limitações da representação de números na máquina se se recorrer a um instrumento de cálculo automático.

Utilizando as notações da página 32, tem-se que:

$$|p_1 - p| \leq \frac{b_1 - a_1}{2} = \frac{b - a}{2}$$

$$|p_2 - p| \leq \frac{b_2 - a_2}{2} = \frac{b_1 - a_1}{2} = \frac{b - a}{2^2}$$

Seguindo este procedimento, iniciado no intervalo $[a, b]$, tem-se que o n -ésimo valor central, p_n , do intervalo $[a_n, b_n]$, é tal que:

$$|p_n - p| \leq \frac{b - a}{2^n} \quad (1)$$

Considerando o exemplo tratado inicialmente na tabela 3.1, suponhamos que se pretende um valor aproximado da solução com erro absoluto inferior a 1 milésima.

Pretende-se determinar o número de iterações necessárias para a aproximação pretendida.

Vejamos então para que valor de n se tem que $\frac{2^{-1}}{2^n} \leq \frac{1}{1000}$.

Resolvendo esta inequação tem-se que $2^n \geq 1000$, ou seja, que $n \geq \log_2 1000$. Como $\log_2 1000 \approx 9,97$, tem-se que qualquer valor no 10º intervalo, e em particular p_{10} , já tem a precisão procurada.

Na tabela 3.3 podemos confirmar o que deduzimos com base na expressão (1)

i	a_i	b_i	p_i	$ p_i - \sqrt{2} $
...
8	1,414062500000	1,421875000000	1,417968750000	0,00375518763
9	1,414062500000	1,417968750000	1,416015625000	0,00180206263
10	1,414062500000	1,416015625000	1,415039062500	0,00082550013
...

Tabela 3.3

Este é um método numérico simples e eficaz para estimar soluções de equações com a precisão que se quiser (sempre limitado pela precisão da aritmética utilizada). Denotando por $Int(x)$ o maior inteiro menor do que x , da equação (1) conclui-se que o número de iterações necessárias para estimar a solução com um erro inferior a ε (assume-se que $\varepsilon < b - a$) é dado, em geral, por

$$n = Int\left(\log_2 \frac{b-a}{\varepsilon}\right) + 1 \quad (2)$$

Como acabamos de ver, usando a expressão (2) é possível conhecer, à priori, o número de iterações a executar para que p_n satisfaça a condição (1). Em geral, isto não é assim. O que é mais comum é que o número de iterações necessárias para cumprir uma certa condição (normalmente designada por critério de paragem) só seja conhecido depois de terminado o cálculo de todas as iterações. A escolha de um critério de paragem depende, em geral, do método utilizado mas também de outros fatores. Falaremos mais deste assunto no contexto de outros métodos que vamos estudar.

3.2.3. ALGUNS CONSTRANGIMENTOS

Até agora usámos o método da bissecção a partir de um intervalo inicial que se sabe conter uma única raiz equação dada $f(x) = 0$. Se $f(a) \times f(b) < 0$ mas o intervalo $[a, b]$ contem mais

do que uma raiz, o número de raízes será ímpar, e o método da bisseção, tal como apresentado antes, será capaz de calcular uma dessas raízes. Um certo número inicial de iterações acabará por descartar, por uma ou mais vezes, intervalos que contêm um número par de raízes, isto é, intervalos em cujos extremos f tem o mesmo sinal.

Como ilustração, consideremos a equação $\frac{1}{4}x^2 \cos(2x) + \sin(2x)$, cuja representação gráfica da figura 3.5 revela cinco das soluções, as quais são em número infinito.

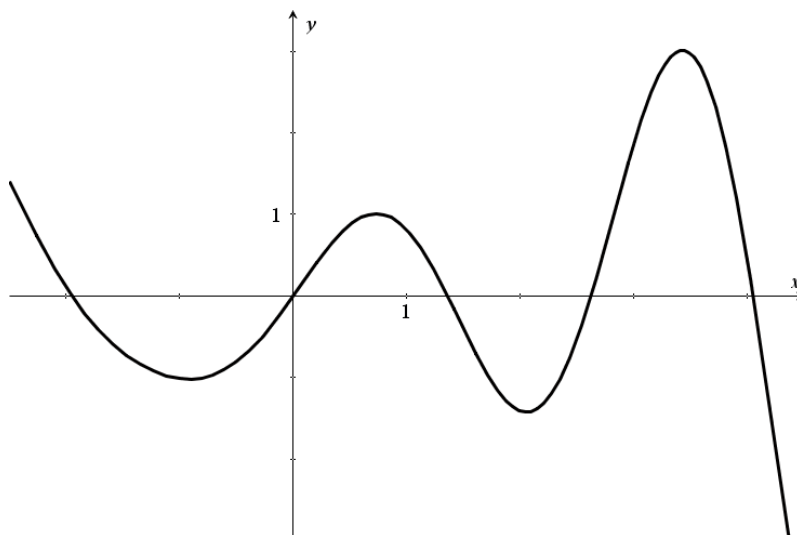


Fig. 3.5

Na tabela 3.4 apresentam-se as primeiras 15 aproximações obtidas pelo método da bisseção (valores centrais) a partir de intervalos iniciais diferentes que contêm um número ímpar de raízes (superior a um). Não é possível saber à priori qual das raízes acabará por ser calculada; como se percebe do exemplo, tal depende do intervalo inicial escolhido.

	$[-2,5]$	$[-2.3; 5]$	$[-2; 4.1]$	$[-2; 4.2]$
i	p_i	p_i	p_i	p_i
1	1,5000000000	1,3500000000	1,0500000000	1,1000000000
2	-0,2500000000	3,1750000000	2,5750000000	2,6500000000
3	-1,1250000000	4,0875000000	1,8125000000	3,4250000000
4	-1,5625000000	3,6312500000	1,4312500000	3,8125000000
5	-1,7812500000	3,8593750000	1,2406250000	4,0062500000
6	-1,8906250000	3,9734375000	1,3359375000	4,1031250000
7	-1,9453125000	4,0304687500	1,3835937500	4,0546875000
8	-1,9726562500	4,0589843750	1,3597656250	4,0304687500
9	-1,9589843750	4,0447265625	1,3478515625	4,0425781250
10	-1,9521484375	4,0518554688	1,3538085938	4,0486328125
11	-1,9487304688	4,0482910156	1,3567871094	4,0456054688
12	-1,9504394531	4,0465087891	1,3552978516	4,0471191406
13	-1,9512939453	4,0473999023	1,3560424805	4,0463623047
14	-1,9508666992	4,0469543457	1,3556701660	4,0467407227
15	-1,9510803223	4,0467315674	1,3554840088	4,0469299316

Tabela 3.4

Fazemos ainda notar que o método da bisseção não pode usar-se para aproximar o valor p da raiz se $f(p) = 0$ é um máximo ou um mínimo local da função f . O problema neste caso é que existe uma vizinhança de p onde o sinal de f é constante. Isto acontece, por exemplo, no caso da simples equação $x^2 = 0$ já que a função é sempre não-negativa. No caso da equação $x^3 = 0$, em que a raiz $p = 0$ é também ponto de inflexão, o método da bisseção não tem qualquer problema em determinar a raiz tripla.

Finalmente, importa referir que a precisão referida anteriormente para as aproximações à solução de uma equação obtidas em número expresso em (1), exemplificado na tabela 3.3, apenas depende do erro absoluto, ou seja, garante casas decimais corretas e não algarismos significativos. Se a precisão se pretende em relação a algarismos significativos e iniciarmos o método num intervalo que contenha uma raiz de módulo inferior a 1, o erro relativo será superior ao erro absoluto. Esta situação não ocorreu nos exemplos referidos porque a solução era superior a 1. É fácil explicar esta situação com base na comparação das expressões para o erro absoluto e para o erro relativo, vistas na página 17. O erro relativo é obtido a partir do erro absoluto pela divisão de um número de módulo inferior a 1.

Observemos a evolução dos erros absolutos e dos erros relativos ao considerar o valor central p_n do intervalo em cada iteração como aproximação da solução nos exemplos apresentados na tabela 3.5.

$x^2 - 2 = 0$ em $[1,2]$			$x^2 - \frac{1}{22222222} = 0$ em $[0,0.001]$		
n	$ p_n - r $	$\left \frac{p_n - r}{r}\right $	n	$ p_n - r $	$\left \frac{p_n - r}{r}\right $
1	0.0857864376269	0.06066017177982	1	0.00043291796064	6.4535599212726
2	0.1642135623731	0.11611652351682	2	0.00018291796064	2.7267799606363
3	0.0392135623731	0.02772817586850	3	0.00005791796064	0.86338998031812
4	0.0232864376269	0.01646599795566	4	0.00000458203936	0.06830500984094
5	0.0079635623731	0.00563108895642	5	0.00002666796064	0.39754248523859
6	0.0076614376269	0.00541745449962	6	0.00001104296064	0.16461873769883
7	0.0001510623731	0.00010681722840	7	0.00000323046064	0.04815686392894
8	0.0037551876269	0.00265531863561	8	0.0000006.7578936	0.01007407295600

Tabela 3.5

É assim claro que um erro relativo na solução (positiva) de $x^2 - 2 = 0$ é sempre inferior ao erro absoluto, o que garante em cada iteração o mesmo número de casas decimais e algarismos significativos corretos e que na equação $x^2 - \frac{1}{22222222} = 0$ já não é assim, pois, por exemplo, a 8ª iteração produz uma aproximação com 6 casas decimais corretas mas apenas um algarismo significativo é garantido.

3.2.4. BREVE CONCLUSÃO

Acabamos de tratar um método numérico de resolução de equações, que apesar de simples e robusto, não converge muito rapidamente para a solução. Como veremos, há métodos que, sob determinadas condições, são muito mais rápidos que o método da bisseção, mas não têm convergência garantida. Um código robusto e eficiente para o cálculo de raízes de equações é sempre a implementação de uma combinação criteriosa de diferentes métodos onde se procura preservar as vantagens principais de cada um deles. Isto acontece certamente com os métodos usados pelas calculadoras. É por esta razão que faz sentido estudar mais do um método de resolução numérica de equações. Do conjunto numeroso de métodos conhecidos, estudaremos mais dois. Um aspeto essencial do nosso estudo será a comparação dos três métodos estudados, com especial destaque para a velocidade de convergência de cada um deles, o que será feito na secção 3.5.

3.3. MÉTODO DO PONTO FIXO

3.3.1 PONTO FIXO

Um ponto fixo para uma função é um número que não se altera pela aplicação da função [22, p. 56], ou seja, p é um ponto fixo de uma função φ se $\varphi(p) = p$. Por exemplo, 0 e 1 são os pontos fixos de φ , definida por $\varphi(x) = x^2$, pois $\varphi(0) = 0$ e $\varphi(1) = 1$.

Note-se que o ponto fixo é a abscissa de um ponto de interseção do gráfico de φ com a reta de equação $y = x$, como se ilustra na figura 3.6.

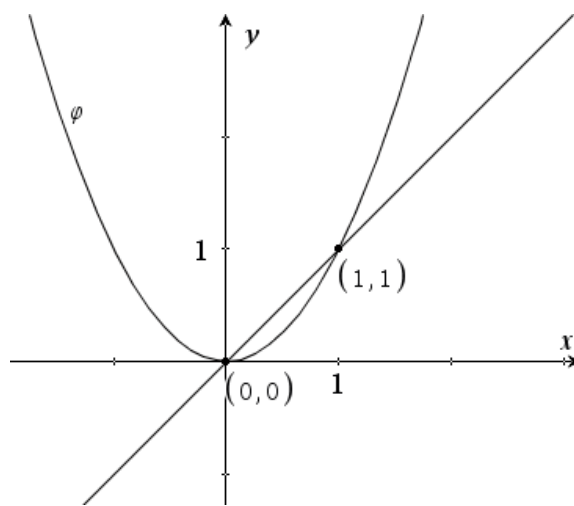


Fig. 3.6

3.3.2 PONTO FIXO E RESOLUÇÃO DE EQUAÇÕES

Podemos procurar pontos fixos de uma função, mas a principal preocupação é resolver equações. Poderá qualquer equação escrita na forma $f(x) = 0$ ser transformada num problema de determinação do ponto fixo de uma função? A resposta é sim. Se p é zero de uma função f , podemos encontrar uma função que dependa de f e que tenha p como ponto fixo. Aliás, os exemplos são em número infinito. Dois deles podem ser

$$\varphi(x) = x - f(x) \text{ ou } \varphi(x) = x + 2f(x).$$

Como obter o ponto fixo da função definida por $\varphi(x) = x^2 - 2$?

Seja p um ponto fixo desta função. Tem-se então que $\varphi(p) = p$, ou seja, $p^2 - 2 = p$, ou de forma equivalente, que $p^2 - p - 2 = 0$. Resolvendo a equação com procedimento analítico, como a utilização da fórmula resolvente, obtemos os pontos fixos -1 e 2 .

É possível garantir a existência de pontos fixos em determinadas funções, e mesmo a unicidade, como se pode concluir pelo teorema 3.2 [22, p. 57].

Teorema 3.2

- i) Se φ é uma função contínua, definida em $[a, b]$ e $\varphi([a, b]) \subset [a, b]$, então φ tem pelo menos um ponto fixo em $[a, b]$.
- ii) Se adicionalmente, $\varphi'(x)$ existe em $]a, b[$ e existe uma constante positiva $k < 1$ tal que $|\varphi'(x)| \leq k$, para todo $x \in]a, b[$, então há exatamente um ponto fixo em $[a, b]$.

Vejamos alguns exemplos.

Na figura 3.7, a função φ representada verifica as condições de existência e também de unicidade, o que se verifica visualmente na figura 3.8.

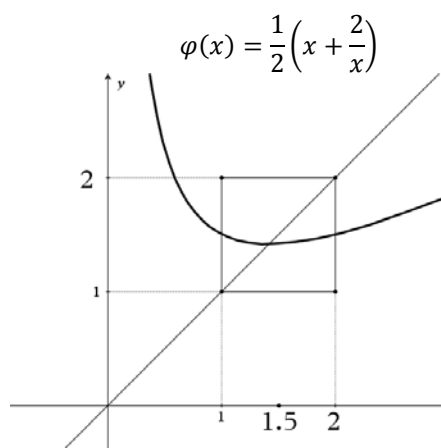


Fig. 3.7

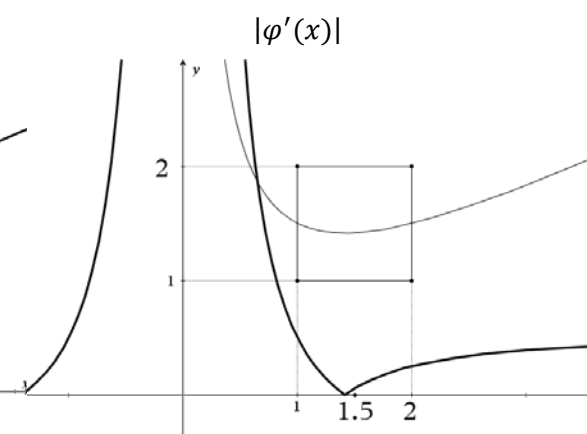


Fig. 3.8

Na figura 3.9 temos uma função φ em que há garantia da existência do ponto fixo, que é único, embora a condição para garantir a unicidade não se verifique porque não há intervalos centrados no ponto fixo onde o valor absoluto da derivada de seja inferior a 1.

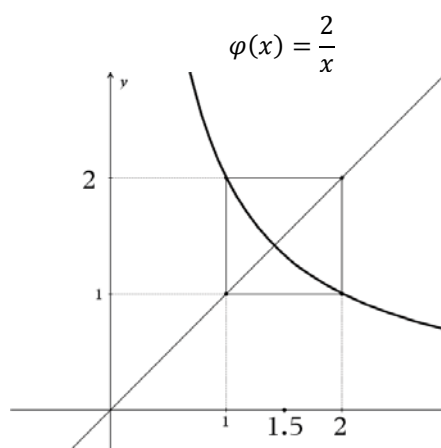


Fig. 3.9

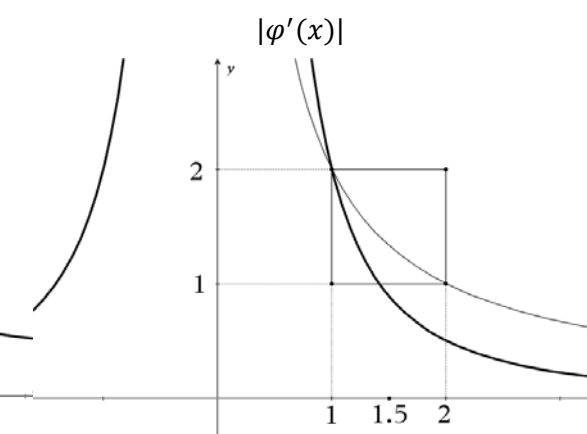


Fig. 3.10

Na figura 3.11, nem se verifica a condição da existência de ponto fixo em $[1,2]$ nem da unicidade ($\varphi'(x) = \frac{3}{2}x^2 \Rightarrow \varphi'(x) > 1, x \in]1,2[$), para a função φ representada, embora exista um único ponto fixo no intervalo referido, o que pode ser rigorosamente concluído com o teorema de Bolzano, uma vez que a função é contínua e $\varphi(1) < 1$ e $\varphi(2) > 2$.

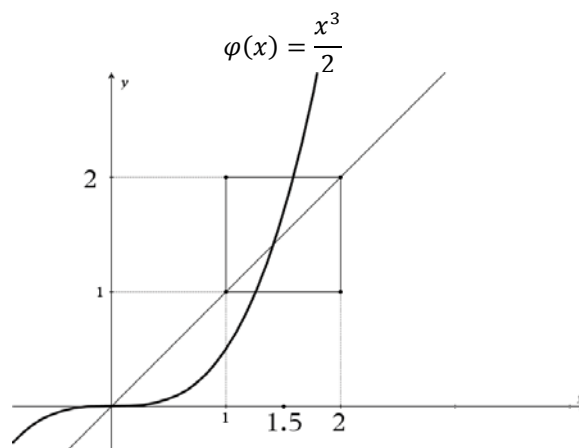


Fig. 3.11

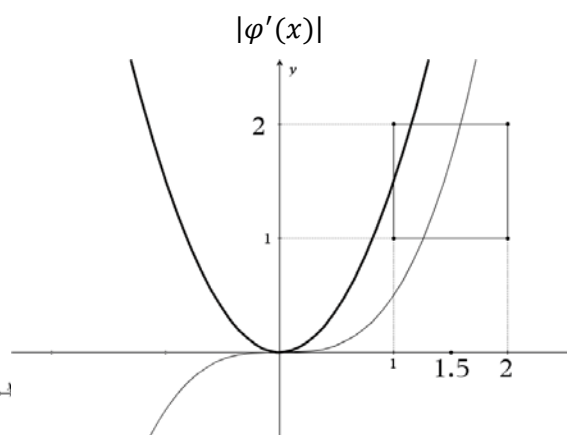


Fig. 3.12

Na figura 3.13, a função φ representada tem mais do que um ponto fixo no intervalo $[1,2]$, verificando-se as condições de existência. A condição de unicidade não é verificada e, neste caso, há mesmo mais do que um ponto fixo.

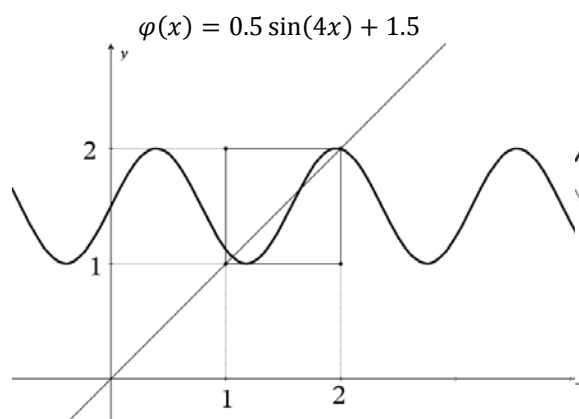


Fig. 3.13

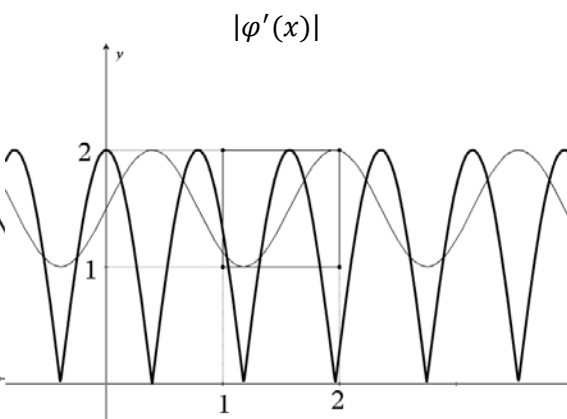


Fig. 3.14

Os pontos fixos podem existir, mas nem sempre podem ser obtidos por procedimentos analíticos. Por exemplo, se $\varphi(x) = e^x - 2$, a equação $e^p - p - 2 = 0$ não pode ser resolvida analiticamente. É sobretudo por causa destas equações, que não se podem resolver por processos analíticos, que assume especial importância um método iterativo de aproximação ao ponto fixo como método iterativo de resolução de equações. Embora as máquinas se possam utilizar para

fazer cálculos não iterativos, os métodos iterativos assumem especial relevância na resolução de equações na máquina.

3.3.3 MÉTODO ITERATIVO BASEADO NA APROXIMAÇÃO AO PONTO FIXO

A figura 3.15, e também a figura 3.16, ilustra o processo iterativo para obter o ponto fixo.

Inicia-se por um valor inicial, x_0 . O valor seguinte é $x_1 = \varphi(x_0)$, e assim sucessivamente:

$$x_0, \quad x_1 = \varphi(x_0), \quad x_2 = \varphi(x_1), \quad \dots, \quad x_{n+1} = \varphi(x_n), \quad \dots$$

Constrói-se, então, uma sucessão de valores que se pretende que seja convergente para o ponto fixo, o qual se obtém no limite. A função φ chama-se **função iteradora**.

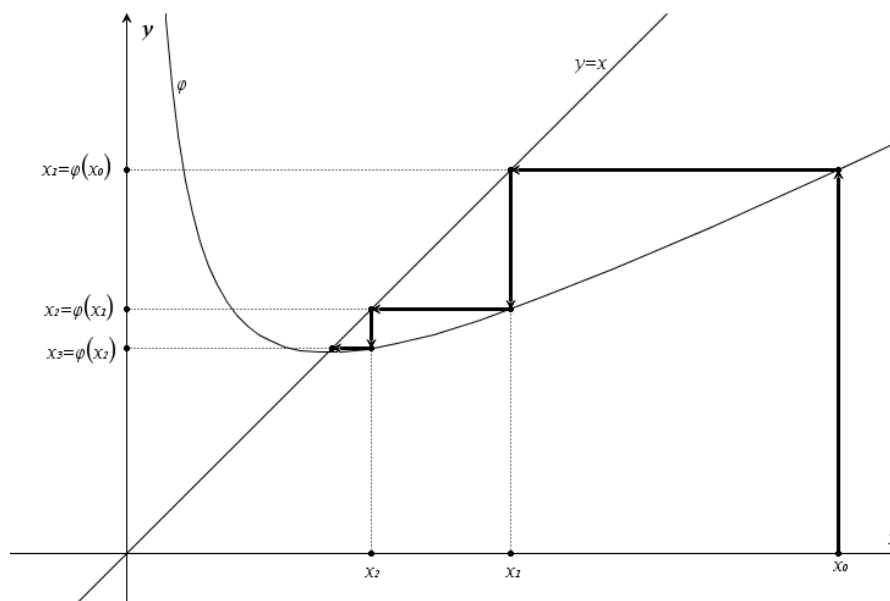


Fig. 3.15

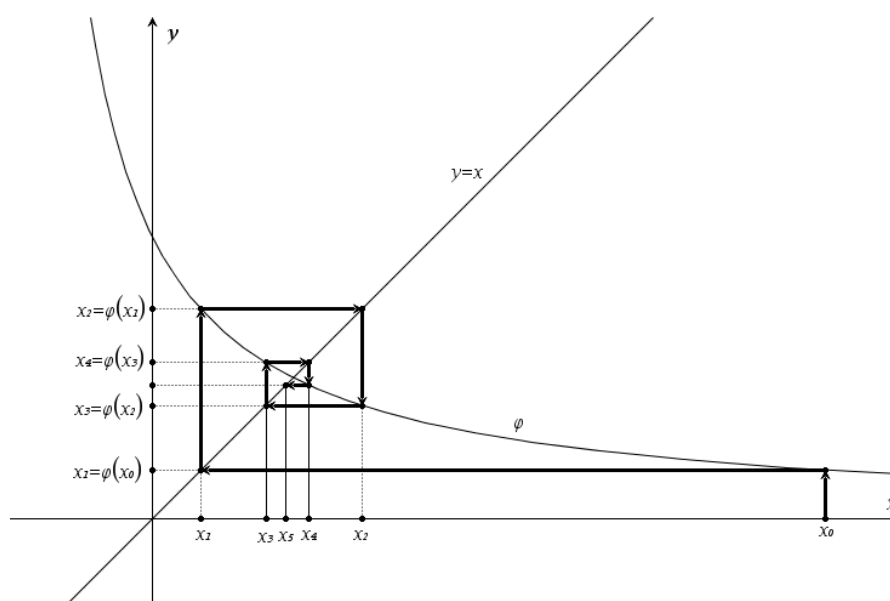


Fig. 3.16

Considere-se, por exemplo, a equação

$$x^2 - 2 = 0. \quad (3)$$

Considerem-se as funções definidas, respetivamente por

$$\varphi(x) = \frac{1}{2} \left(x + \frac{2}{x} \right), \quad (4)$$

por

$$\varphi(x) = \frac{2}{x} \quad (5)$$

e por

$$\varphi(x) = \frac{x^3}{2} \quad (6)$$

as quais estão representadas graficamente nas figuras 3.7 a 3.12. Note-se que uma solução da equação (3) é ponto fixo de qualquer uma das funções (4), (5) e (6). Reciprocamente, e porque para $x \neq 0$ as três equações $x = \varphi(x)$ são equivalentes e equivalentes à equação (3), um ponto fixo não nulo de cada uma das três funções é solução de (3). A partir de valores iniciais x_0 , podemos definir as sucessões $x_{n+1} = \varphi(x_n)$.

Com $x_0 = 1.5$, obtemos os valores registados na tabela 3.6. Note-se que quando nada for dito, os valores colocados em tabelas são obtidos com a calculadora TI-Nspire Cx.

n	x_n , de (4)	x_n , de (5)	x_n , de (6)
0	1,5	1,5	1,5
1	1,41666666666667	1,33333333333333	1,6875
2	1,41421568627451	1,50000000000000	2,4027099609375
3	1,41421356237469	1,33333333333333	6,93544051044864
4	1,41421356237309	1,50000000000000	166,79850621909
5	1,41421356237309	1,33333333333333	2320312,47606181
6	1,41421356237309	1,50000000000000	6246107146539600000
7	1,41421356237309	1,33333333333333	Overflow
8	1,41421356237309	1,50000000000000	Overflow
9	1,41421356237309	1,33333333333333	Overflow
10	1,41421356237309	1,50000000000000	Overflow

Tabela 3.6

Observa-se que com 10 iterações, a sucessão (x_n) que resulta de iterações da função (4), parece estabilizar em torno da solução positiva da equação (3), a sucessão que resulta de iterações da função (5) oscila entre dois valores e a sucessão que resulta de iterações da função (6) toma valores superiores ao maior valor representável na máquina.

Nas figuras 3.7 a 3.12, cujas funções já foram anteriormente analisadas relativamente ao teorema 3.2, e conjugando com a análise da tabela 3.6, conclui-se que pelo facto de existir ponto fixo de φ isso não garante a convergência dos valores iterados.

Este comportamento diferenciado das sucessões obtidas com diferentes funções iteradoras sugere a existência de condições, sejam elas relativas à função iteradora utilizada ou ao ponto inicial, que determinam a convergência (ou divergência) e o ponto fixo para o qual a sequência converge. Por exemplo, com $x_0 = -1$, x_k em $\varphi(x) = \frac{1}{2}\left(x + \frac{2}{x}\right)$ converge, não para a solução positiva da equação (3), como aconteceu com $x_0 = 1.5$, mas para a solução negativa..

Observando as mesmas figuras, nomeadamente as representações das funções módulo das funções derivadas de φ , parece que a condição referida para a unicidade poderá ter alguma relação com condições para a convergência das sucessões das funções iteradoras, pois talvez não seja coincidência o facto de que quando não se verifica a condição de unicidade do ponto fixo também não se tem convergência e que quando esta é verificada se tem convergência.

Um exemplo referido em [16, p. 34] permite reforçar geometricamente estas ideias sobre condições de convergência das sucessões das funções iteradoras.

Vejamos uma situação simples.

As figuras 3.17 e 3.18 apresentam iterações para duas funções lineares, $\varphi_1(x) = -\frac{3}{2}x + \frac{5}{2}$ e $\varphi_2(x) = -\frac{1}{2}x + \frac{3}{2}$.

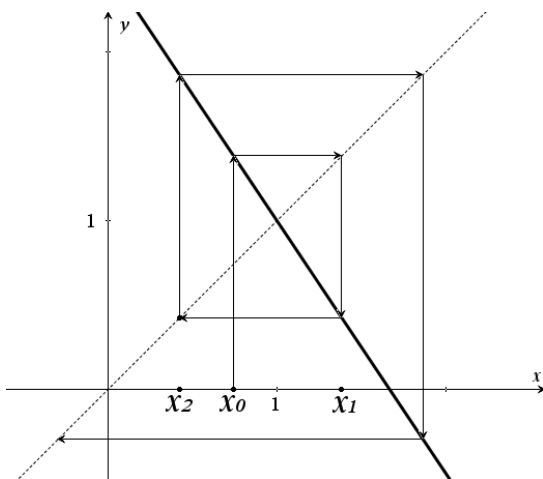


Fig. 3.17

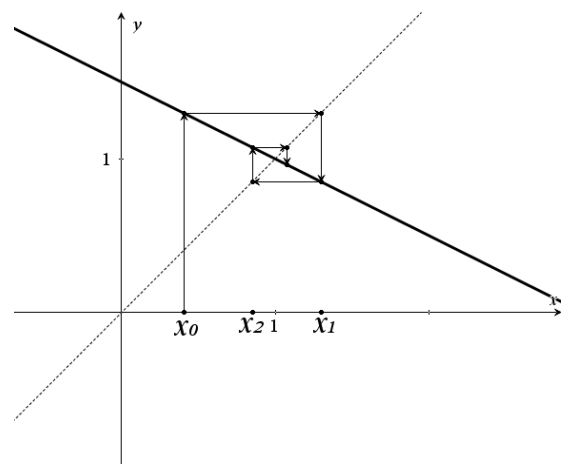


Fig. 3.18

Em cada caso o ponto fixo é $x = 1$, mas $\varphi_1'(1) = \left|-\frac{3}{2}\right| > 1$ e $\varphi_2'(1) = \left|-\frac{1}{2}\right| < 1$.

Observando as sucessivas setas que ilustram geometricamente o desenvolvimento do método iterativo pode encontrar-se a razão desta diferença. Uma vez que o módulo do declive da reta de $\varphi_1(x)$ é superior a 1 no ponto fixo, os segmentos orientados verticais, os que representam a mudança de x_n para x_{n+1} , crescem à medida que se efetuam as iterações, o que tem como

consequência uma “espiral” que “abre”, revelando a divergência da sucessão, mesmo que o valor inicial esteja muito próximo de 1. O contrário sucede com $\varphi_2(x)$, pelo que a “espiral” “fecha”, revelando a convergência.

Assim, aparentemente $|\varphi'(x)|$ faz a diferença entre a divergência e a convergência.

3.3.4 CONDIÇÕES DE CONVERGÊNCIA DO MÉTODO DO PONTO FIXO

Vejam uma condição suficiente para a convergência de uma sucessão $x_{n+1} = \varphi(x_n)$ para um ponto fixo p , em que é φ uma função iteradora.

Note-se que se esta sucessão convergir, apenas pode ser convergente para p , pois no limite $x_{n+1} = x_n$, pelo que a convergência não pode ser para outro valor que não seja um ponto fixo.

Para cada valor iterado, considere-se o erro de x_n em relação ao ponto fixo p , $e_n = x_n - p$.

Teorema 3.3

Seja φ uma função contínua num intervalo I centrado no ponto fixo p , e diferenciável no interior de I .

A sucessão $x_{n+1} = \varphi(x_n)$ é convergente para o ponto fixo p em I se $|\varphi'(x)| < 1$.

Demonstração:

Note-se que $e_{n+1} = x_{n+1} - p = \varphi(x_n) - \varphi(p)$

Considerando intervalos de extremos p e x_n tem-se que, pelo teorema do valor médio, há um valor θ entre p e x_n tal que

$$\varphi(x_n) - \varphi(p) = (x_n - p) \times \varphi'(\theta) \quad (7)$$

Assim, tem-se que $e_{n+1} = e_n \times \varphi'(\theta)$ e, por ser $|\varphi'(\theta)| < 1$, a sucessão dos erros tende para zero, isto é, x_n converge para p .

Manipulando algebricamente a expressão (7) pode obter-se sucessivamente:

$$\frac{x_{n+1} - p}{x_n - p} = \varphi'(\theta)$$

$$\frac{(x_{n+1} - x_n) + (x_n - p)}{x_n - p} = \varphi'(\theta)$$

$$\frac{x_{n+1} - x_n}{x_n - p} + 1 = \varphi'(\theta)$$

$$(x_n - p)(\varphi'(\theta) - 1) = x_{n+1} - x_n$$

$$e_n = x_n - p = \frac{1}{\varphi'(\theta) - 1} (x_{n+1} - x_n) \quad (8)$$

A expressão (8) permite concluir que considerando as condições do teorema 3.3, a diferença entre duas iterações sucessivas é uma boa estimativa do erro absoluto se $\varphi'(x) \approx 0$ numa vizinhança de p , a qual piora à medida que φ' se aproxima de 1. [26]

Na figura 3.19 estão reunidas as funções derivadas das funções iteradoras observadas anteriormente, e já apresentadas nas figuras 3.8 (gráfico c1), 3.10 (gráfico c2) e 3.12 (gráfico c3). Estes exemplos ilustram a relação próxima entre a condição de unicidade de ponto fixo e a condição de convergência, permitindo, assim, o teorema 3.3 explicar os diferentes

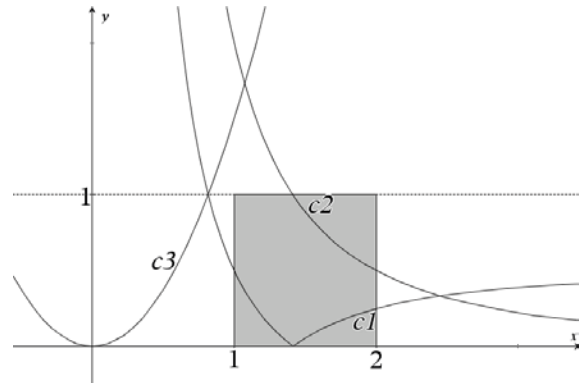


Fig. 3.19

comportamentos que tiveram estas três funções iteradoras, expressos nos resultados apresentados na tabela 3.6.

Note-se que, se considerarmos $I = [1,2]$, a única função derivada da função iteradora que toma, em módulo, valores inferiores a 1 é a que tem o gráfico da função módulo representado por c1, enquanto das restantes funções derivadas, uma tem o módulo sempre superior a 1 (c3) e a outra apenas para alguns valores tem o módulo inferior a 1, a qual, mesmo reduzindo o intervalo, centrado em $\sqrt{2}$ nunca toma, em módulo, valores inferiores a 1.

Façamos agora uma comparação entre a função iteradora

$$\varphi(x) = \frac{1}{2} \left(x + \frac{2}{x} \right)$$

e duas novas funções iteradoras para a mesma equação (3), $x^2 - 2 = 0$,

$$\varphi_3(x) = 2 - \frac{2}{x+2}$$

e

$$\varphi_4(x) = \frac{4x+2}{x+4}$$

cujas representações gráficas e respetivos módulos da função derivada se podem observar nas figuras 3.20 e 3.21.

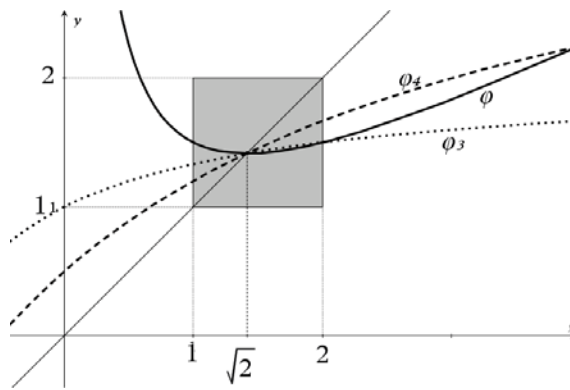


Fig. 3.20.

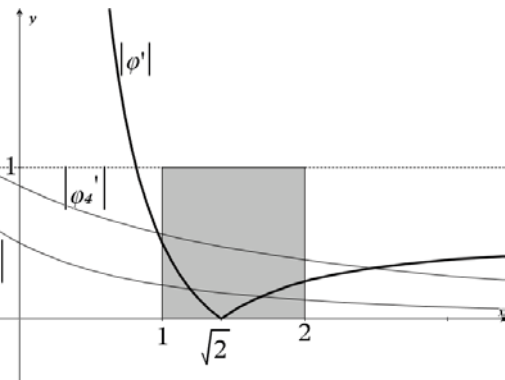


Fig. 3.21.

Observando o comportamento das funções iteradoras e das respectivas funções derivadas, pode dizer-se que são boas funções no sentido da convergência para o ponto fixo.

Observe-se na tabela 3.7 as primeiras 10 iterações de cada uma das funções para um valor inicial $x_0 = 1.5$, por exemplo.

n	x_n , com (φ)	x_n , com (φ_3)	x_n , com (φ_4)
1	1,41666666666667	1,42857142857143	1,45454545454545
2	1,41421568627451	1,41666666666667	1,43333333333333
3	1,41421356237469	1,41463414634146	1,42331288343558
4	1,41421356237309	1,41428571428571	1,41855203619909
5	1,41421356237309	1,41422594142259	1,41628392484342
6	1,41421356237309	1,41421568627451	1,41520197348134
7	1,41421356237309	1,41421392677674	1,41468553369587
8	1,41421356237309	1,41421362489487	1,41443895257125
9	1,41421356237309	1,41421357310014	1,41432120250399
10	1,41421356237309	1,41421356421356	1,41426496944338

Tabela 3.7

Podemos confirmar nestes resultados o que já se esperava da figura 3.21: a sequência obtida com a função iteradora φ converge mais rapidamente do que as sequências obtidas com as funções iteradoras φ_3 e φ_4 porque o módulo da função derivada toma valores mais próximos de zero na vizinhança do ponto fixo.

Esta análise do valor da função derivada da função iteradora vai ter um papel de relevo no método iterativo que será mais à frente abordado, um caso particular deste.

Relativamente ao zero $-\sqrt{2}$ da equação (3), da análise dos módulos das derivadas dessas funções iteradoras, próximo de $-\sqrt{2}$ (figura 3.23), podemos perceber que, com exceção de φ , não é possível associar sucessões de valores iterados convergentes para $-\sqrt{2}$

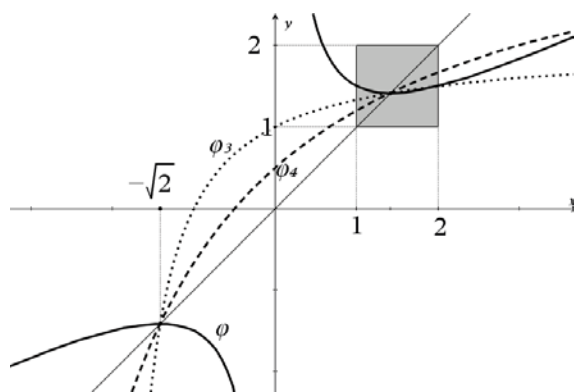


Fig. 3.22

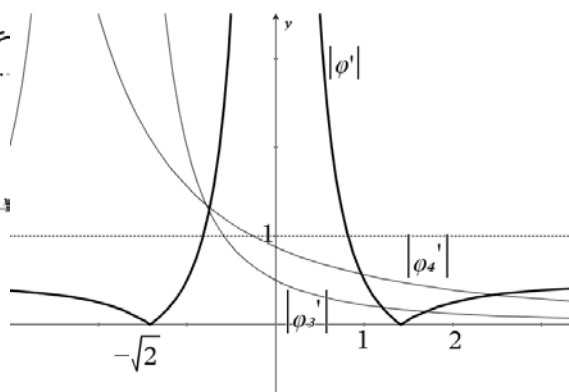


Fig. 3.23

As tabelas 3.8 e 3.9 ilustram este facto, em que se podem observar duas opções diferentes para o valor inicial.

Valor inicial = -1			
k	x_k, com (φ)	x_k, com (φ_3)	x_k, com (φ_4)
1	-1,500000000000000	0,000000000000000	-0,666666666666667
2	-1,416666666666667	1,000000000000000	-0,200000000000000
3	-1,41421568627451	1,333333333333333	0,31578947368421
4	-1,41421356237469	1,400000000000000	0,75609756097561
5	-1,41421356237309	1,41176470588235	1,05641025641026
...			
21	-1,41421356237309	1,41421356237309	1,41421056114785
22	-1,41421356237309	1,41421356237309	1,41421212901038
23	-1,41421356237309	1,41421356237309	1,41421287780999
24	-1,41421356237309	1,41421356237309	1,41421323543102
25	-1,41421356237309	1,41421356237309	1,41421340622808

Tabela 3.8

Valor inicial = -1,4142			
k	x_k, com (φ)	x_k, com (φ_3)	x_k, com (φ_4)
1	-1,41421356243813	-1,41413451689997	-1,41418516513265
2	-1,41421356237309	-1,41375291375291	-1,41415410385260
3	-1,41421356237309	-1,41153081510932	-1,41408906882591
4	-1,41421356237309	-1,39864864864853	-1,41395290581162
5	-1,41421356237309	-1,32584269662858	-1,41366784520753
...			
16	-1,41421356237309	1,41421322976698	-0,29537403465982
17	-1,41421356237309	1,41421350530690	0,22094102590073
18	-1,41421356237309	1,41421355258208	0,68320407366685
19	-1,41421356237309	1,41421356069322	1,01059364918123
20	-1,41421356237309	1,41421356208487	1,20591990087089
...			
31	-1,41421356237309	1,41421356237309	1,41414726177701
32	-1,41421356237309	1,41421356237309	1,41418189733447
33	-1,41421356237309	1,41421356237309	1,41419843930760
34	-1,41421356237309	1,41421356237309	1,41420633969404
35	-1,41421356237309	1,41421356237309	1,41421011287295

Tabela 3.9

Para as funções iteradoras φ_3 e φ_4 verifica-se a convergência para $\sqrt{2}$, isto porque há valores iterados que “caiem” em intervalos em torno de $\sqrt{2}$, onde os módulos das funções derivadas são inferiores a 1.

Escolher boas funções iteradoras pode não ser uma tarefa fácil, a qual, em certa medida, pode ser enquadrada com resultados teóricos, mas que não deixa de fazer apelo à criatividade de quem utiliza o método. Importa aprofundar um pouco mais estas questões.

3.3.5 O MÉTODO DO PONTO FIXO NA RESOLUÇÃO DE EQUAÇÕES DO 2º GRAU

Temos vindo a usar o exemplo de uma equação polinomial do 2º grau que não tem parte linear, $x^2 - 2 = 0$. Estudamos uma função iteradora que não só permitiu uma rápida convergência, como foi capaz de produzir as duas soluções, a partir de valores iniciais diferentes, claro.

Para $c > 0$, a equação

$$x^2 - c = 0, \quad (9)$$

pode ser reescrita na forma equivalente, com $x \neq 0$,

$$x = \varphi(x), \text{ com } \varphi(x) = \frac{1}{2} \left(x + \frac{c}{x} \right). \quad (10)$$

Pelo teorema 3.3, temos que a sequência das iterações

$$x_0, \quad x_1 = \varphi(x_0), \quad x_2 = \varphi(x_1), \quad \dots, \quad x_{n+1} = \varphi(x_n), \quad \dots$$

converge para uma raiz da equação (10), e consequentemente raiz da equação (9), se se verificar a condição:

$$|\varphi'(x)| < 1,$$

num intervalo que contenha a raiz e a estimativa inicial x_0 .

Aplicando este resultado à equação (10), tem-se que:

$$\varphi'(x) = \frac{1}{2} \left(1 - \frac{c}{x^2} \right) = \frac{(x - \sqrt{c})(x + \sqrt{c})}{2x^2}. \quad (11)$$

Note-se que as raízes da equação (9) são simétricas. Seja r a raiz positiva.

Tem-se que r e $-r$ são as raízes e que $r^2 = c$.

Considerando a equação (11), tem-se que na vizinhança de r , ou de $-r$,

$$|\varphi'(x)| \approx \frac{0}{2r^2} = 0$$

Assim sendo, esta função iteradora φ produz, de um modo geral, um método sempre convergente para a raiz mais próxima da estimativa considerada, o que já fazia supor a análise particular com $c = 2$.

Seguindo um processo de raciocínio análogo para a função iteradora $\varphi(x) = c + \frac{c-c^2}{x+c}$, cuja concretização para $c = 2$ foi vista nas tabelas 3.8, 3.9 e 3.10:

$$\varphi'(x) = \frac{c^2-c}{(x+c)^2} = \frac{r^4-r^2}{(x+c)^2} = \frac{r^2(r-1)(r+1)}{(x+r^2)^2}. \quad (12)$$

Numa vizinhança da raiz positiva r de (9), da equação (12), tem-se que:

$$|\varphi'(x)| \approx \left| \frac{r^2(r-1)(r+1)}{r^2(1+r)(1+r)} \right| = \left| \frac{r-1}{r+1} \right|. \quad (13)$$

Pode concluir-se que $\left| \frac{r-1}{r+1} \right| < 1 \Leftrightarrow r > 0$, o que significa que com esta função iteradora a convergência dá-se sempre para a raiz positiva, o que vem confirmar o que se observou nas tabelas 3.8 e 3.9.

Nota: A relação (13) é também válida para $r < 0$ e a conclusão é a mesma, tratando-se de uma raiz negativa.

Considere-se agora, de um modo mais geral, uma equação polinomial do 2º grau com duas soluções distintas, a qual se pode escrever:

$$x^2 - sx + p = 0, \quad (14)$$

onde

$$s = r_1 + r_2$$

é a soma das raízes r_1 e r_2 e

$$p = r_1 \times r_2$$

é o produto das mesmas raízes.

A equação (14) pode ser reescrita na forma equivalente:

$$x = \sigma(x) = s - \frac{p}{x}, \quad (15)$$

$$x = \omega(x) = -\frac{p}{x-s}, \quad (16)$$

com $x \neq 0$ e $x \neq s$, respetivamente.

A teoria da convergência do método do ponto fixo diz-nos que as sucessões das iterações:

$$x_0, \quad x_1 = \sigma(x_0), \quad x_2 = \sigma(x_1), \quad \dots, \quad x_{n+1} = \sigma(x_n), \quad \dots$$

e

$$x_0, \quad x_1 = \omega(x_0), \quad x_2 = \omega(x_1), \quad \dots, \quad x_{n+1} = \omega(x_n), \quad \dots$$

(x_0 é a estimativa inicial, não necessariamente a mesma para cada uma das sucessões)

convergem para uma raiz da equação (15) e da equação (16), respetivamente, e portanto também para uma raiz da equação (14), se se verificarem as condições:

$$|\sigma'(x)| < 1. \quad (17)$$

$$|\omega'(x)| < 1, \quad (18)$$

respetivamente num intervalo que contenha a raiz e a estimativa inicial x_0 .

Aplicando este resultado à equação (15), vem que $\sigma'(x) = \frac{p}{x^2}$ e na vizinhança de uma raiz, por exemplo r_2 , é

$$|\sigma'(x)| \approx \left| \frac{r_1 \times r_2}{r_2^2} \right| = \left| \frac{r_1}{r_2} \right| \text{ (converge se } |r_2| > |r_1| \text{)}$$

o que mostra que esta escolha para a função iteradora produz um método convergente para a raiz de maior valor absoluto mas não para a de raiz de menor valor absoluto.

Aplicando o mesmo resultado à equação (16), vem que $\omega'(x) = \frac{p}{(x-s)^2}$ e na vizinhança de uma raiz, por exemplo r_2 , é

$$|\omega'(x)| \approx \frac{|r_1 \times r_2|}{[r_2 - (r_1 + r_2)]^2} = \frac{|r_1 \times r_2|}{r_1^2} = \frac{|r_2|}{|r_1|} \text{ (converge se } |r_2| < |r_1| \text{)}$$

o que mostra que agora a convergência é para a raiz de menor valor absoluto.

Vejamos a ilustração destes resultados para a equação geral do 2º grau com um exemplo.

Exemplo 3.2

Considere-se a equação:

$$x^2 - 2x - 5 = 0. \quad (19)$$

Considerem-se, a partir das funções (15) e (16), as funções iteradoras:

$$\sigma(x) = 2 + \frac{5}{x}, \quad (20)$$

$$\omega(x) = \frac{5}{x-2}. \quad (21)$$

Observe-se, nas tabelas 3.10 e 3.11, o comportamento dos valores iterados aplicando o método do ponto fixo a partir de diferentes valores iniciais.

Para melhor análise importa fazer uma referência às soluções da equação,

$$1 - \sqrt{6} = -1.449489742783178 \dots$$

$$1 + \sqrt{6} = +3.449489742783178 \dots$$

σ		
k	$x_0 = -1,5$	$x_0 = 3,5$
1	-1,3333333333333333	3,42857142857143
2	-1,7500000000000000	3,4583333333333333
3	-0,85714285714286	3,44578313253012
4	-3,8333333333333333	3,45104895104895
5	0,69565217391304	3,44883485309017
.....		
31	3,44948974176328	3,44948974278307
32	3,44948974321174	3,44948974278322
33	3,44948974260309	3,44948974278316
34	3,44948974285885	3,44948974278319
35	3,44948974275138	3,44948974278317

Tabela 3.10

ω		
k	$x_0 = 1,5$	$x_0 = 3,5$
1	-1,42857142857143	3,3333333333333333
2	-1,4583333333333333	3,7500000000000000
3	-1,44578313253012	2,85714285714286
4	-1,45104895104895	5,8333333333333333
5	-1,44883485309017	1,30434782608696
.....		
31	-1,44948974278307	-1,44948974176328
32	-1,44948974278322	-1,44948974321174
33	-1,44948974278316	-1,44948974260309
34	-1,44948974278319	-1,44948974285885
35	-1,44948974278317	-1,44948974275138

Tabela 3.11

3.4. MÉTODO DE NEWTON-RAPHSON

3.4.1. INTRODUÇÃO E NOTA HISTÓRICA

Um dos mais conhecidos e poderosos métodos de resolução de equações é o método de Newton-Raphson, largamente utilizado devido à sua relativa simplicidade e à sua habitual rapidez de convergência. Em [20] Atkinson considera-o o melhor método para encontrar soluções de uma equação. Como qualquer método numérico, este também não é infalível, pois pode não ser possível a sua utilização e nem sempre converge rapidamente.

O método de Newton-Raphson, também conhecido pelo método das tangentes, aparece na obra de Isaac Newton, *De Methodus Fluxionum et Serierum infinitorum*, escrita entre 1664 e 1671, para encontrar a solução da equação

$$y^3 - 2y - 5 = 0, \quad (22)$$

e foi melhorado por Joseph Raphson, como é referido em [27]. Este método, tal como hoje o conhecemos, envolve derivadas, mas nem Newton nem Raphson terão usado explicitamente as “fluxões” (derivadas na terminologia introduzida por Newton). No processo de resolução da equação (22), Newton (ver, por exemplo, [27]) começa por considerar um valor inicial igual a 2, que difere da solução menos que a sua décima parte. Substituiu, na expressão (22), y por $2 + p$, dando origem à equação

$$p^3 + 6p^2 + 10p - 1 = 0. \quad (23)$$

Para determinar um valor para p , desprezou $p^3 + 6p^2$ e de $10p - 1 = 0$ obteve $p = 0.1$, que se aproxima muito da solução de (23), e será portanto 2.1 uma solução aproximada da equação (22).

Em seguida considera $p = 0.1 + q$ e substituindo p na equação (23) obtém

$$q^3 + 6.3q^2 + 11.23q + 0.061 = 0. \quad (24)$$

Voltando a desprezar os termos não lineares, obtém $q = -0.0054 \dots$, que é a solução da equação $11.23q + 0.061 = 0$ e aproxima a solução de (24). Newton volta a substituir a variável fazendo $q = -0.0054 + r$ e, desprezando novamente os termos não lineares, chega a $r \approx -0.00004852$.

Assim obtém Newton a solução $2 + p + q + r$, que é aproximadamente 2.09455148 para a equação (22).

Note-se que o erro relativo desta estimativa, obtida com três iterações apenas, é inferior a 10^{-9} , o que diz bem da rapidez de convergência.

Este procedimento, usado por Newton, é equivalente ao que se baseia na fórmula iterativa que usa derivadas e que apresentaremos na secção seguinte. A versão de Raphson será mais simples e abrangente que a de Newton, e mais próxima da formulação atual [27]. Tal justifica a associação dos nomes dos dois cientistas britânicos na designação pela qual é conhecido o método atualmente.

Para perceber o funcionamento deste método, começaremos pela interpretação geométrica. Logo, passaremos a uma abordagem analítica mais profunda baseada nos desenvolvimentos em série de potências.

3.4.2. INTERPRETAÇÃO GEOMÉTRICA E ALGORITMO

Seja f uma função real de variável real, contínua e derivável, que tem pelo menos uma raiz real, r . Considere-se uma estimativa desta raiz, x_0 .

Note-se que uma equação da reta tangente ao gráfico de f no ponto de abcissa x_0 é

$$y = f'(x_0)(x - x_0) + f(x_0). \quad (25)$$

Seja x_1 a abcissa do ponto de interseção da reta com o eixo das abcissas. Ao substituir na equação anterior y por 0 e x por x_1 , obtém-se $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$.

Se se considerar agora x_1 como nova estimativa de r e se for possível continuar a repetir o processo, obtém-se uma sucessão (x_n) , que poderá ser convergente para r , definida por

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (26)$$

A figura 3.24 faz-se a interpretação geométrica do método de Newton-Raphson.

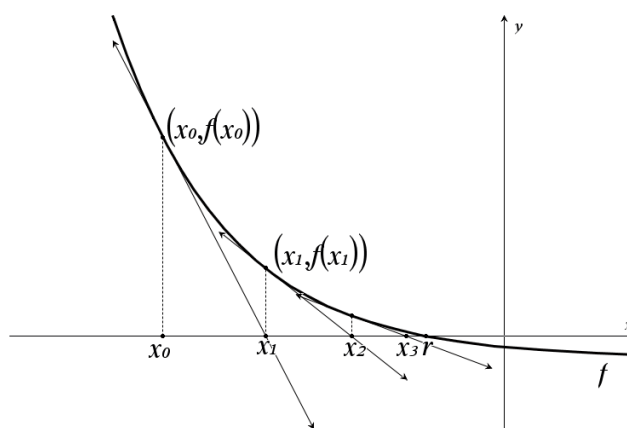


Fig. 3.24

Dada uma estimativa inicial x_0 , considera-se a reta tangente ao gráfico de f no ponto $(x_0, f(x_0))$, a qual, se não for horizontal, vai encontrar o eixo das abcissas em x_1 . De seguida, considera-se a reta tangente ao gráfico de f no ponto $(x_1, f(x_1))$, a qual, não sendo também horizontal, vai encontrar o eixo das abcissas em x_2 . Iterando este procedimento, espera-se que as sucessivas aproximações x_n estejam cada vez mais próximas de r , solução da equação do tipo $f(x) = 0$.

Vejamos a aplicação deste método no caso particular da equação (22). A função f , definida por $f(x) = x^3 - 2x - 5$ está representada graficamente na figura 3.25 para valores no intervalo $[1.995; 2.110]$. Note-se que $f'(x) = 3x^2 - 2$.

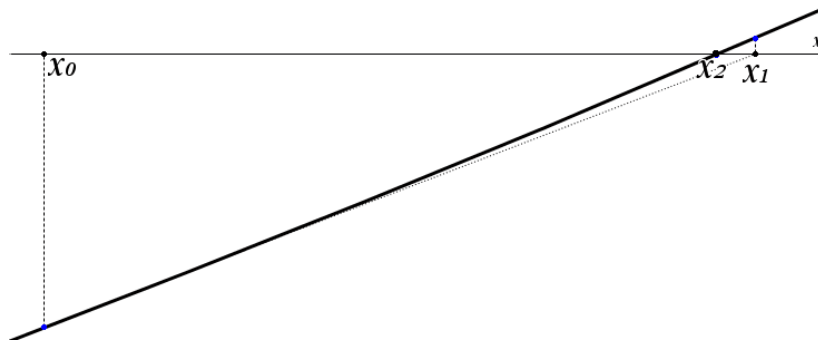


Fig. 3.25

Uma equação da reta tangente ao gráfico de f no ponto de abscissa $x_0 = 2$ é

$$y = 10x - 21, \quad (27)$$

pelo que a primeira iteração do método leva à aproximação $x_1 = \frac{21}{10} = 2,1$.

A reta tangente ao gráfico de f no ponto de abscissa x_1 tem a equação

$$y = 11.23x - 23.522 \quad (28)$$

E dá $x_2 = \frac{23.522}{11.23} \approx 2.0945681$. Continuando, vem $x_3 \approx 2.0945514$, e assim sucessivamente.

Note-se que esta 3ª aproximação da solução é, com o mesmo número de algarismos significativos, a que foi obtida por Newton, referida na secção anterior. A parte linear das equações que aí se foram obtendo estão explicitamente relacionadas com as equações (27) e (28), o que revela bem da estreita relação do procedimento de Newton com este que acabamos de usar e que é usualmente implementado pela fórmula (26).

De maneira equivalente, a fórmula (26) pode obter-se a partir da razão trigonométrica $\tan \theta = f'(x_n) = \frac{f(x_n)}{x_n - x_{n+1}}$ que é válida para funções com diferentes comportamentos na vizinhança de x_n , como se ilustra nas figuras 3.26 - 3.29.

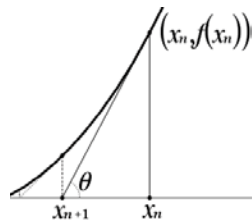


Fig. 3.26

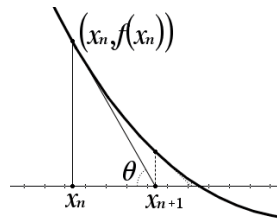


Fig. 3.27

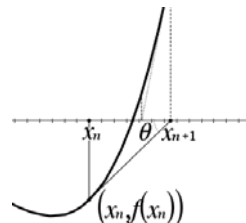


Fig. 3.28

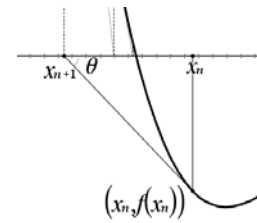


Fig. 3.29

Vamos de seguida fazer um tratamento mais analítico que nos levará de novo à fórmula (26) mas também permitirá compreender melhor o funcionamento do método de Newton-Raphson.

3.4.3. O MÉTODO DE NEWTON-RAPHSON E A FÓRMULA DE TAYLOR

Se f é uma função real de variável real com derivadas contínuas até à ordem n num ponto x_0 do seu domínio, o **polinómio de Taylor** da função f no ponto x_0 é definido por

$$p_n(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!}f''(x_0) + \dots + \frac{(x - x_0)^n}{n!}f^{(n)}(x_0) \quad (29)$$

Note-se que $p_n(x_0) = f(x_0)$, $p_n'(x_0) = f'(x_0)$, ... e $p_n^{(n)}(x_0) = f^{(n)}(x_0)$, isto é, a função f e o respetivo polinómio de Taylor de grau n , bem como as primeiras n derivadas de f e p_n , coincidem no ponto x_0 ; isto sugere que p_n e f têm um comportamento muito semelhante, ou seja, que o polinómio de Taylor aproxima bem f na vizinhança de x_0 .

A partir de (29), com $n = 1$, obtém-se o polinómio de Taylor de grau 1

$$p_n(x) = f(x_0) + (x - x_0)f'(x_0) \quad (30)$$

Que está afinal de acordo com a formulação anterior feita em termos de reta tangente à curva $y = f(x)$ no ponto de abscissa x_0 .

Nas figuras 3.30 e 3.31 podem observar-se as representações gráficas de duas funções e dos respetivos polinómios de Taylor de grau 1 nos pontos $x = 0$ e $x = 1$, respetivamente, podendo

comprovar-se que tais polinómios correspondem às retas tangentes aos gráficos das funções nesses pontos.

$$f(x) = \sin(x) ; p_1(x) = x$$

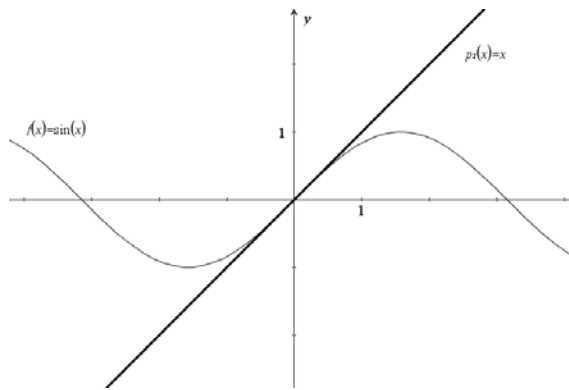


Fig. 3.30

$$f(x) = \ln(x) ; p_1(x) = x - 1$$

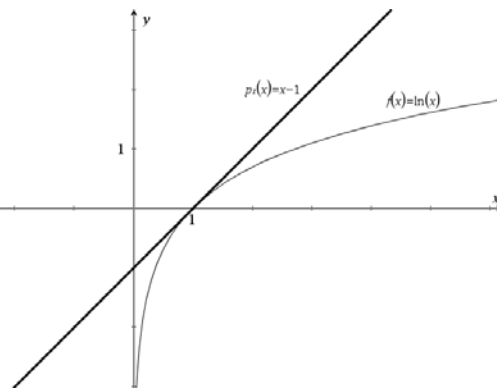


Fig. 3.31

Nas figuras 3.32 e 3.33 podem observar-se representações gráficas das mesmas funções e dos respetivos polinómios de Taylor de grau 3 e 2, respetivamente, das funções em $x = 0$ e $x = 1$. Como expetável, quanto maior for o grau do polinómio de Taylor melhor ele se ajusta à função na vizinhança do ponto considerado.

$$f(x) = \sin(x) ; p_3(x) = x - \frac{x^3}{6}$$

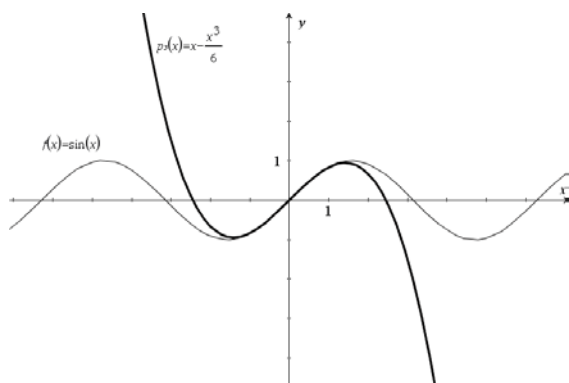


Fig. 3.32

$$f(x) = \ln(x) ; p_3(x) = x - 1 - \frac{(x-1)^2}{2}$$

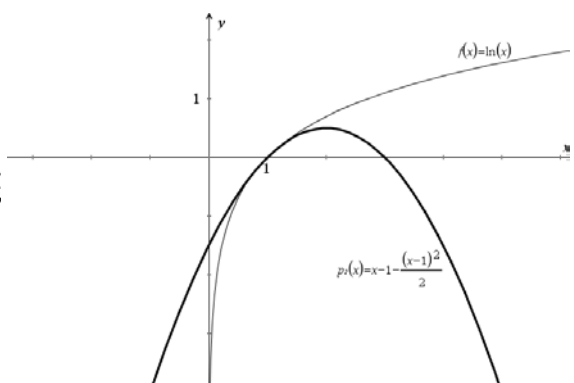


Fig. 3.33

O teorema seguinte constitui a base para a dedução analítica da fórmula iterativa (26) bem como para a expressão do erro cometido em cada iteração.

Teorema 3.4 (Teorema de Taylor)

Seja f uma função com derivadas contínuas até à ordem $k + 1$ num intervalo $[a, b]$ (futuramente usaremos a notação $f \in C^{k+1}[a, b]$ para referir este facto) e seja $x_0 \in]a, b[$. Então, para $x \in [a, b]$,

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \dots + \frac{(x - x_0)^k}{k!} f^{(k)}(x_0) + R_k(x) \quad (31)$$

onde

$$R_k(x) = \frac{(x - x_0)^{k+1}}{(k+1)!} f^{(k+1)}(\theta) \quad (32)$$

sendo θ um valor entre x e x_0 . R_k é conhecido por **resto de Lagrange**.

A demonstração pode ser encontrada, por exemplo, em [13, p. 28].

Em [22, p. 67] supõe-se que temos uma função $f \in C^2[a, b]$. Se x_n , nesse intervalo, é uma aproximação de r tal que $f'(x_n) \neq 0$ e $|r - x_n|$ é “pequeno”, ou seja, x_n é suficientemente próximo de r , a fórmula (31) dá para $k = 1$ e $x = r$.

$$f(r) = f(x_n) + (r - x_n)f'(x_n) + \frac{(r - x_n)^2}{2} f''(\theta) \quad (33)$$

para algum θ entre r e x_0 .

Se r é zero da função f , ou seja, $f(r) = 0$, de (33) tem-se que

$$0 = f(x_n) + (r - x_n)f'(x_n) + \frac{(r - x_n)^2}{2} f''(\theta) \quad (34)$$

Dividindo ambos os membros desta equação por $f'(x_n) \neq 0$ e rearranjando a expressão obtida resulta

$$r = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f''(\theta)}{2f'(x_n)} (r - x_n)^2 \quad (35)$$

(faz-se notar que θ depende de x_n). Sendo $|r - x_n|$ “suficientemente” pequeno (em qualquer caso bastante menor que a unidade), o último termo de (35), que envolve $(r - x_n)^2$ pode desprezar-se e temos

$$r \approx x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (36)$$

Partindo de um valor inicial x_0 conveniente, podemos assim construir uma sucessão de aproximações que, sob certas condições converge para a solução r .

3.4.4. ALGUNS CONSTRANGIMENTOS

Dissemos antes que a convergência do método de Newton-Raphson depende de algumas condições. Por exemplo, em certos casos é necessário que x_0 seja escolhido bastante próximo da solução (para garantir, em particular que a função derivada não se anula no intervalo definido por x_0 e r). O comportamento do método é tanto melhor quanto mais afastada de zero é a função derivada nos valores iterados. Apresentamos de seguida alguns exemplos ilustrativos do funcionamento do método e das dificuldades que podem surgir. Alguns deles serão retomados mais tarde.

Exemplo 3.3

Para resolver a equação $3x - x^3 = \frac{15}{32}$, aplicando o Método de Newton-Raphson, escreve-se a equação na forma $f(x) = 0$ com $f(x) = x^3 - 3x + \frac{15}{32}$.

Sendo $f'(x) = 3x^2 - 3$, a fórmula (26) dá neste caso

$$x_{n+1} = \frac{2x_n^3 - \frac{15}{32}}{3x_n^2 - 3} \quad (37)$$

Na tabela 3.12 podemos observar o comportamento dos primeiros 5 termos da sucessão para diferentes valores de x_0 .

	$x_0 = -1$	$x_0 = -0.8$	$x_0 = -0.75$
i	x_i		
1	n.d.	1.3821759259259	1
2	n.d.	1.7619514612585	n.d.
3	n.d.	1.6585484897334	n.d.
4	n.d.	1.6479986058614	n.d.
5	n.d.	1.6478912539222	n.d.

(n.d. – não definido)

Tabela 3.12

Uma vez que o denominador da expressão no segundo membro de (37), isto é, a derivada $f'(x) = 3x_n - 3$, se anula para $x = \pm 1$, o método falha se alguma das aproximações assumir

um destes valores. Tal acontece obviamente no caso em que tomámos $x_0 = -1$, mas também acontece com $x_0 = -0.75$ porque neste caso tem-se

$$x_1 = \frac{2(-0.75)^3 - \frac{15}{2}}{3(-0.75)^2 - 3} = 1$$

Relativamente a $x_0 = -0.8$, o método poderá levar, aparentemente, a uma solução muito próxima de 0.1647, o que pode ser comprovado com mais iterações. Considerando doze dígitos, a partir de x_7 obtém-se sempre 1.6478912428518.

Nas figuras 3.34 a 3.36 pode observar-se a interpretação gráfica do comportamento do método para as diferentes estimativas iniciais consideradas.

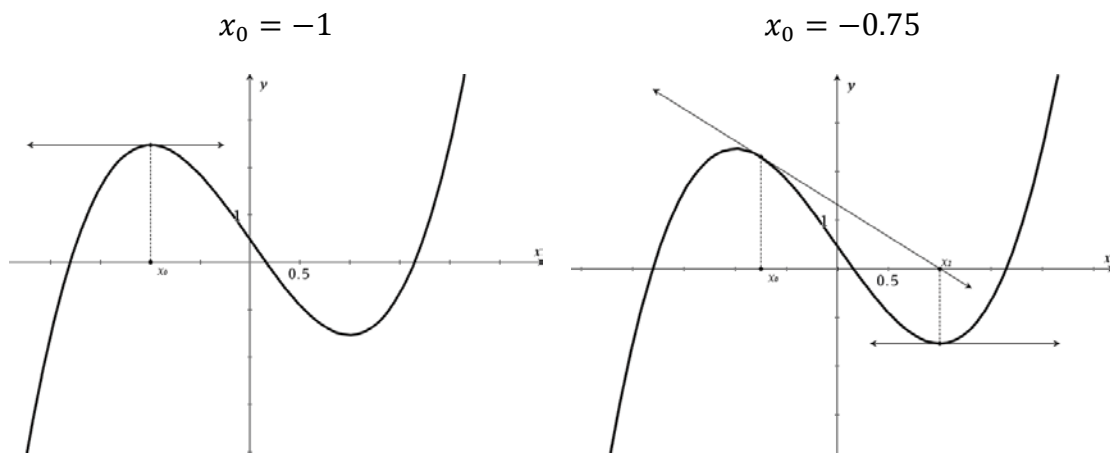


Fig. 3.34

Fig. 3.35

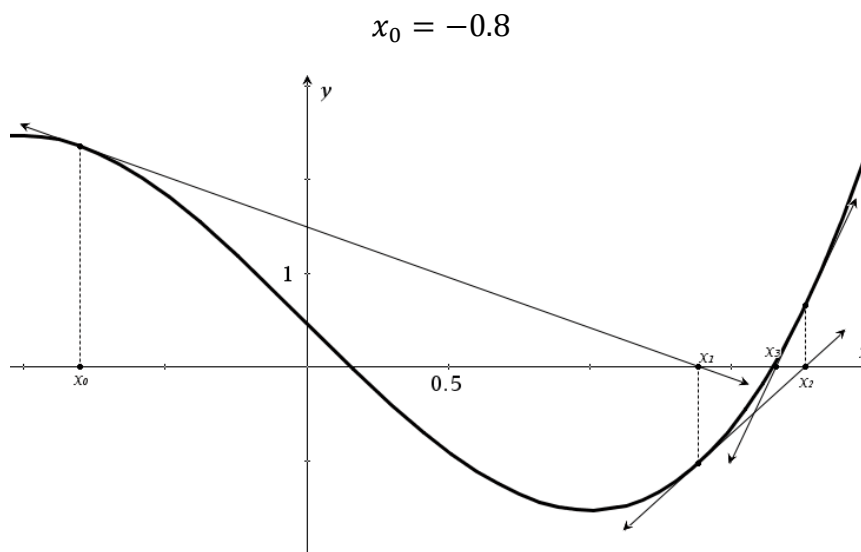


Fig. 3.36

Exemplo 3.4 [16, p. 58]

No caso da equação $4x^4 - 6x^2 - 2.75 = 0$, a fórmula iterativa é

$$x_{n+1} = x_n - \frac{4x_n^4 - 6x_n^2 - \frac{11}{4}}{16x_n^3 - 12x_n}$$

Começando com $x_0 = 0.5$ produz uma sucessão de valores que são alternadamente -0.5 e 0.5 , ou seja, não converge. Obviamente, o mesmo acontece se se começar com $x_0 = -0.5$. Embora também no caso desta função existam pontos onde se anula a derivada, com $x_0 = \pm 0.5$ o método falha por razão diferente (no gráfico da figura 3.37 ilustra-se o problema).

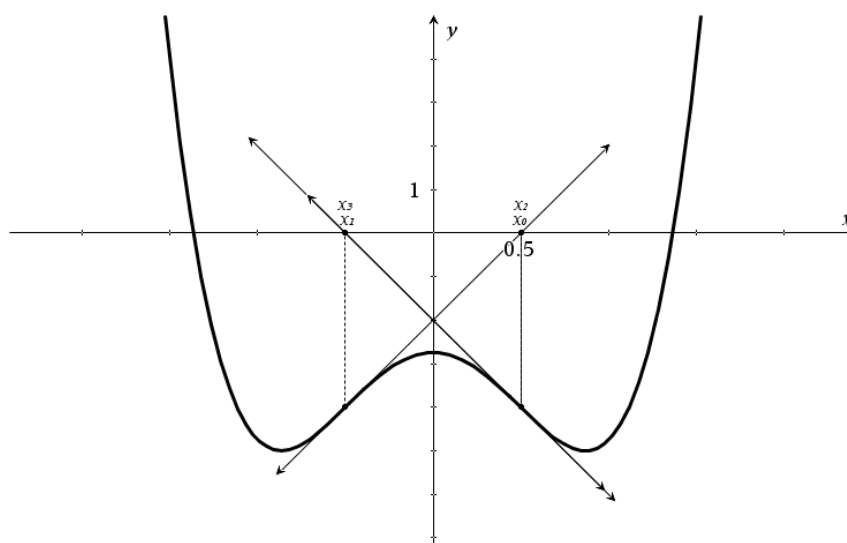


Fig. 3.37

Exemplo 3.5 [28]

Para determinar a solução da equação $\sqrt[3]{x} = 0$, que é obviamente 0, a fórmula (26) com a função $f(x) = \sqrt[3]{x}$ é

$$x_{n+1} = x_n - \frac{x_n^{\frac{1}{3}}}{\frac{1}{3}x_n^{-\frac{2}{3}}} = x_n - 3x_n = -2x_n$$

Partindo de $x_0 = 1$, obtém-se a sucessão $-2, 4, -8, 16, -32, \dots$ que é, como se vê, divergente e como tal o método volta a não ser eficaz. Aliás, neste exemplo, o método diverge seja qual for a aproximação inicial x_0 , como facilmente se conclui da fórmula $x_{n+1} = -2x_n$.

Exemplo 3.6

Neste exemplo pretende-se resolver a equação $\cos(x) = 0$, aplicando mais uma vez o método de Newton-Raphson. Nas figuras 3.38 e 3.39 pode observar-se o efeito da aplicação do método, na primeira com $x_0 = 0.364$ e na seguinte com $x_0 = 0.43$. Na tabela 3.13 podem observar-se as primeiras oito iterações para cada uma das situações.

$$f(x) = \cos(x) ; x_0 = 0.364$$

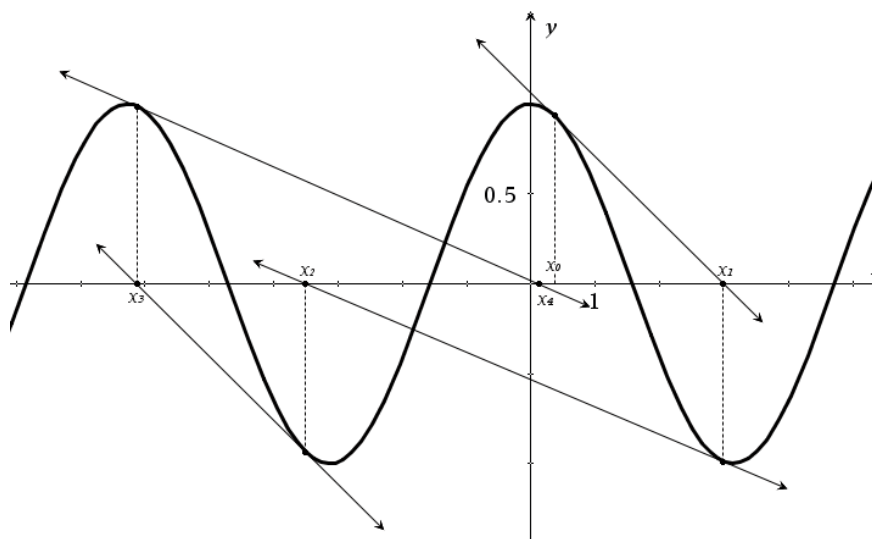


Fig. 3.38

$$f(x) = \cos(x) ; x_0 = 0.43$$

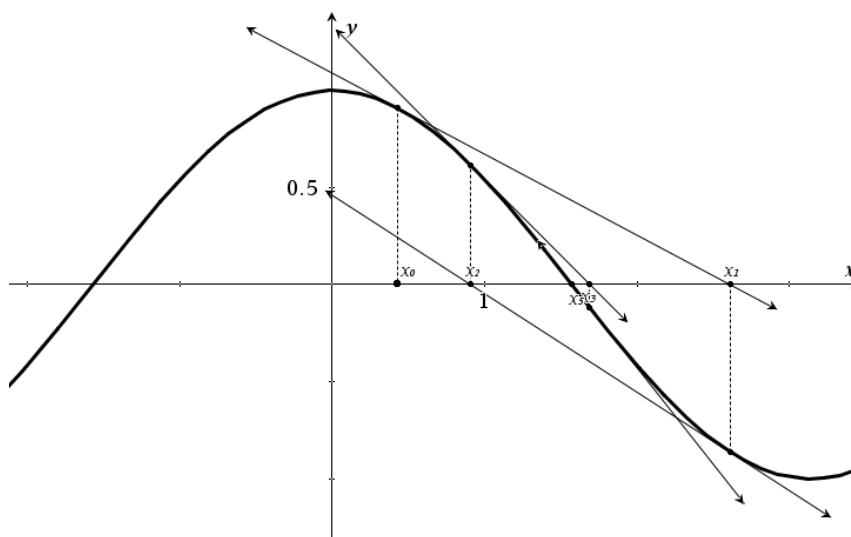


Fig. 3.39

	$x_0 = 0.364$	$x_0 = 0.43$
i	x_i	
1	2,98883396288064	2,61044954066851
2	-3,50643947692387	0,90818739389429
3	-6,12460702889125	1,68848119242774
4	0,12847862319052	1,57024999888757
5	7,86900119891287	1,57079632684925
6	7,85398050446471	1,57079632679490
7	7,85398163397448	1,57079632679490
8	7,85398163397448	1,57079632679490

Tabela 3.13

Observa-se um efeito curioso quando se considera $x_0 = 0.364$, designado na literatura inglesa por “root jumping”, em que as iterações “saltam” de forma aparentemente desordenada entre diferentes soluções, sendo difícil prever aquela para a qual acabará por tender. Esta situação não significa que o método não convirja para uma determinada raiz mas não é fácil determinar *a priori* para qual delas. No entanto, ao observar a figura 3.38, podemos suspeitar que possa ocorrer um comportamento oscilatório que tentamos comprovar em experiências numéricas. Na tabela 3.14 podem observar-se as primeiras quatro iterações para $x_0 = 0.36400290378$.

i	x_i
1	2,98881395672732
2	-3,50559555804613
3	-6,13040660565738
4	0,36400270792083

Tabela 3.14

Exemplo 3.7

A equação $(2x - 1)^5 = 0$ tem a solução única $x = 0.5$ mas como $x_0 = 0.2$ a fórmula

$$x_{n+1} = x_n - \frac{(2x_n - 1)^5}{10(2x_n - 1)^4} = \frac{4}{5}x_n + \frac{1}{10}$$

produz uma sequência cujas primeiras iterações são apresentadas na tabela 3.15. Embora a sequência pareça convergir para a raiz, tal acontece muito lentamente.

$$x_0 = 0.2$$

i	x_i	i	x_i
1	0.26	6	0.4213568
2	0.308	7	0.43708544
3	0.3464	8	0.449668352
4	0.37712	9	0.4597346816
5	0.401696	10	0.46778774528

Tabela 3.15

3.4.5. CONVERGÊNCIA DO MÉTODO

Ficou bem ilustrado nos exemplos anteriores que nem sempre o método de Newton-Raphson converge e mesmo quando converge não é garantido que a convergência seja rápida. Vamos agora analisar com mais atenção a convergência do método, o que nos permitirá melhorar a compreensão do seu funcionamento e perceber também porque é, em geral, a convergência tão rápida, o que o torna o método tão importante e tão utilizado.

Teorema 3.5

Seja $f \in C^2[a, b]$. Se $r \in]a, b[$ é tal que $f(r) = 0$ e $f'(r) \neq 0$, então existe $\delta > 0$ tal que qualquer que seja a aproximação inicial $x_0 \in [r - \delta, r + \delta]$, o método de Newton-Raphson gera uma sucessão (x_n) convergente para r .

Este teorema alerta para a necessidade de valores iniciais “suficientemente” próximos da solução, o que é usual na maioria dos métodos iterativos de resolução de equações. Na prática, a utilidade deste teorema é muito limitada pelo facto de não ser conhecido o valor δ em cada caso.

A demonstração do teorema 3.5 apresentada em [22, p. 70] baseia-se na análise da iteração funcional $x_{n+1} = \varphi(x_n)$, $n \geq 0$, com

$$\varphi(x) = x - \frac{f(x)}{f'(x)}. \quad (38)$$

Isto é relevante para uma análise mais aprofundada das questões de convergência, e confirma o facto de que o método de Newton-Raphson pode ser interpretado como um caso especial do método do ponto fixo [16, p. 53], o qual foi tratado anteriormente. Por isso importa ter em consideração o teorema 3.3 que dá a condição suficiente de convergência $|\varphi'(x)| < 1$.

Note-se que uma vez que $f(r) = 0$ e $f'(r) \neq 0$, de (38) vem que $\varphi(r) = r$, donde, procurar a solução r de $f(x) = 0$ é o mesmo que procurar o ponto fixo r da função $\varphi(x)$.

Utilizando (38) resulta

$$\varphi'(x) = \frac{f(x) \times f''(x)}{[f'(x)]^2} \quad (39)$$

e portanto $\varphi'(r) = 0$. Sendo φ' contínua, existe um intervalo $I =]r - \sigma, r + \sigma[$ onde se verifica a condição $|\varphi'(x)| < 1$. Se a aproximação inicial pertencer a I , então o teorema 3.3 garante a convergência das iterações $\varphi(x_n)$ para o ponto fixo r , ou seja, das iterações

$$x_n - \frac{f(x_n)}{f'(x_n)}$$

para a solução r da equação $f(x) = 0$. Podemos por isso afirmar, como em [16, p. 53], que se $\varphi'(r) \neq 0$, o método de Newton-Raphson é localmente convergente.

Para continuar a analisar as questões de convergência é importante a introdução da definição de ordem de convergência.

Definição 3.1 [22, p. 79]

Seja (x_n) uma sucessão convergente para r , com $x_n \neq r, \forall n \in \mathbb{N}$. Se há constantes positivas c e α tais que $\lim \frac{|e_{n+1}|}{|e_n|^\alpha} = c$, então diz-se que (x_n) converge para r com ordem α e constante de erro assintótico c .

Dois ordens de convergência merecem especial atenção, pelo que importa denominar o tipo de convergência nesses casos.

- i) Se $\alpha = 1$ (e $c < 1$), a sucessão (x_n) tem **convergência linear**;
- ii) Se $\alpha = 2$, a sucessão (x_n) tem **convergência quadrática**.

No geral, uma sucessão com ordem superior de convergência converge mais rapidamente do que uma de ordem inferior, o que mostra que interessa encontrar métodos que produzam convergência de ordens superiores, como se ilustra no exemplo 3.8.

Exemplo 3.8 [22, p. 79].

Suponhamos que se têm duas sucessões (a_n) e (b_n) , que convergem para 0, a primeira com convergência linear e a segunda com convergência quadrática, e suponhamos ainda que $c = 0.5$.

Assim, tem-se que $\lim \frac{|a_{n+1}|}{|a_n|} = 0.5$ e que $\lim \frac{|b_{n+1}|}{b_n^2} = 0.5$.

Para n “suficientemente” grande, podemos escrever

$$\frac{|a_{n+1}|}{|a_n|} \approx 0.5 \quad , \quad \frac{|b_{n+1}|}{b_n^2} \approx 0.5$$

No caso da convergência linear tem-se que

$$|a_n| \approx 0.5|a_{n-1}| \approx 0.5^2|a_{n-2}| \approx \dots \approx 0.5^n|a_0|$$

e no caso da convergência quadrática é

$$|b_n| \approx 0.5|b_{n-1}|^2 \approx 0.5^2|b_{n-2}|^2 \approx \dots \approx 0.5^{2^{n-1}}|b_0|^{2^n}$$

A expressão final diz bem da maior rapidez da convergência quadrática, o que pode ser comprovado na tabela 3.16, em que se apresentam os primeiros oito termos das sucessões de termos gerais $0.5^n|a_0|$ e $0.5^{2^{n-1}}|b_0|^{2^n}$ com $a_0 = b_0 = 1$.

	convergência linear	Convergência quadrática
i	x_i	
1	5.00000×10^{-1}	5.00000×10^{-1}
2	2.50000×10^{-1}	1.25000×10^{-1}
3	1.25000×10^{-1}	7.81250×10^{-3}
4	6.25000×10^{-2}	3.05176×10^{-5}
5	3.12500×10^{-2}	4.65661×10^{-10}
6	1.56250×10^{-2}	1.08420×10^{-19}
7	7.81250×10^{-3}	5.87747×10^{-39}
8	3.90625×10^{-3}	1.72723×10^{-77}

Tabela 3.16

O teorema seguinte mostra que a convergência do método de Newton-Raphson é quadrática desde que r seja uma raiz simples da equação.

Teorema 3.6

Seja f uma função contínua e derivável, com pelo menos uma raiz r . A partir do valor inicial x_0 , considere-se a sucessão de valores iterados definida por:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, n \geq 0 \wedge f'(x_i) \neq 0, i = 1, \dots, n$$

Se $f''(x)$ é contínua e $f'(r) \neq 0$ (r de multiplicidade 1), então

$$\lim \frac{|e_{n+1}|}{|e_n|^2} = c \text{ com } c = \frac{1}{2} \frac{f''(r)}{f'(r)}.$$

Demonstração:

De (35) e (36) vem

$$e_{n+1} = \frac{1}{2} \frac{f''(\theta)}{f'(x_n)} e_n^2$$

e quando n cresce tanto x_n como θ (que está entre r e x_n) convergem para r e tomando limites na expressão anterior a conclusão é imediata.

Retomando a equação usada por Newton e já referida na secção 3.4.1 pode considerar-se a função f é definida por $f(x) = x^3 - 2x - 5$, a qual respeita os pressupostos do teorema 3.6.

Considere-se a fórmula iteradora do método de Newton-Raphson

$$x_{n+1} = x_n - \frac{x_n^3 - 2x_n - 5}{3x_n^2 - 2} = \frac{2x_n^3 + 5}{3x_n^2 - 2}, n = 0, 1, 2, \dots$$

Na tabela 3.17 podem observar-se as primeiras oito iterações e o comportamento das sucessões definidas por $|e_n| = |x_n - r|$ e por $\frac{|e_{n+1}|}{e_n^2}$.

	x_n	$ e_n $	$\frac{ e_{n+1} }{e_n^2}$
i	x_i		
0	1.000000000000000	1.09455148×10^0	4.09455
1	7.000000000000000	4.90544852×10^0	0.11100
2	4.76551724137931	2.67096576×10^0	0.17580
3	3.34870275948028	1.25415128×10^0	0.27786
4	2.53159964100251	$4.37048159 \times 10^{-1}$	0.41550
5	2.17391588493923	$7.93644034 \times 10^{-2}$	0.52903
6	2.09788368644176	$3.33220490 \times 10^{-3}$	0.56147
7	2.09455771585006	$6.23430773 \times 10^{-6}$	0.56289
8	2.09455148156421	$2.18776108 \times 10^{-11}$	

Tabela 3.17

Uma vez que a solução é, com arredondamento para o mais próximo com quinze algarismos, 2.09455148154233, observa-se que com apenas oito iterações temos onze algarismos corretos. A partir da quarta iteração, o número correto de algarismos corretos aproximadamente duplica a cada nova iteração, o que é característico da convergência quadrática, confirmando-se assim a rapidez de convergência.

De acordo com o que é referido no teorema 3.6, a tabela 3.17 aparentemente mostra-nos um valor de c muito próximo de 0.563. Na realidade, calculando $\frac{1}{2} \frac{f''(r)}{f'(r)}$, ou seja, $\frac{3r}{3r^2 - 2}$, em que r é a raiz real do polinómio $x^3 - 2x - 5$, tem-se que $c = 0.5629789$, com 7 dígitos significativos, o que ilustra a previsível adequação da prática à teoria.

Na figura 3.61 podem observar-se os procedimentos de cálculo, na calculadora TI-Nspire, para obter o valor de c , com interesse particular por se observar o valor exato de r , obtido em [29].

$$\frac{3 \cdot r}{3 \cdot r^2 - 2} \rightarrow r$$

$$\frac{\sqrt[3]{\frac{\sqrt{643} + 5}{6 \cdot \sqrt{3}} + \frac{5}{2}} + \frac{2}{2}}{\sqrt[3]{\frac{\sqrt{643} + 5}{6 \cdot \sqrt{3}} + \frac{5}{2}}} \rightarrow r$$

2.09455148154

0.562978945778

Fig. 3.40

É necessário especial atenção aos pressupostos do teorema 3.6 para haver garantia de convergência quadrática do método de Newton-Raphson. Mesmo que seja convergente, nem sempre a convergência é quadrática.

Já vimos que tomar uma estimativa inicial distante da solução pode não pôr em causa a convergência mas o número de iterações necessárias (e o conseqüente esforço de cálculo) dependerá, em geral, do ponto de partida. Como se pode ver pelos resultados apresentados na tabela 3.17, se o nosso ponto de partida fosse x_4 , que coincide com o valor correto em apenas um algoritmo, apenas quatro iterações seriam necessárias para obter a raiz com onze algoritmos corretos. Noutros casos, como ficou ilustrado em vários dos exemplos apresentados, a convergência depende da qualidade da aproximação inicial.

O PROBLEMA DAS RAIZES MÚLTIPLAS

A não verificação do pressuposto de que a raiz seja de multiplicidade 1, ou seja, que $f'(r) \neq 0$ coloca definitivamente em causa a velocidade de convergência do método de Newton-Raphson, como se viu no exemplo 3.7, o qual será mais desenvolvido de seguida. Da sucessão iteradora utilizada para resolver a equação $(2x - 1)^5 = 0$ vem

$$x_{n+1} - \frac{1}{2} = \frac{4}{5}x_n + \frac{1}{10} - \frac{1}{2} = \frac{4}{5}\left(x_n - \frac{1}{2}\right)$$

e

$$\frac{e_{n+1}}{e_n} = \frac{4}{5}$$

o que mostra que a convergência é apenas linear, de acordo com a Definição 3.1.

Vejamos mais um exemplo, adaptado de [16, p. 56], do comportamento do método de Newton-Raphson para raízes de multiplicidade superior a um.

Exemplo 3.9

Seja $f(x) = \sin x + x^2 \cos x - x^2 - x$.

Vamos iterar até obter um valor da solução da equação $f(x) = 0$ com quatro casas decimais corretas, iniciando com $x_0 = 1$.

De

$$f'(x) = \cos x + 2x \cos x - x^2 \sin x - 2x - 1$$

$$f''(x) = -\sin x + 2 \cos x - 4x \sin x - x^2 \cos x - 2$$

$$f'''(x) = -\cos x - 6 \sin x - 6x \cos x + x^2 \sin x$$

tem-se $f(0) = f'(0) = f''(0) = 0$ e $f'''(0) = -1$, pelo que $x = 0$ é uma raiz de f de multiplicidade 3. Na figura 3.41 estão representadas geometricamente as primeiras quatro iterações, as quais se podem observar numericamente na tabela 3.18.

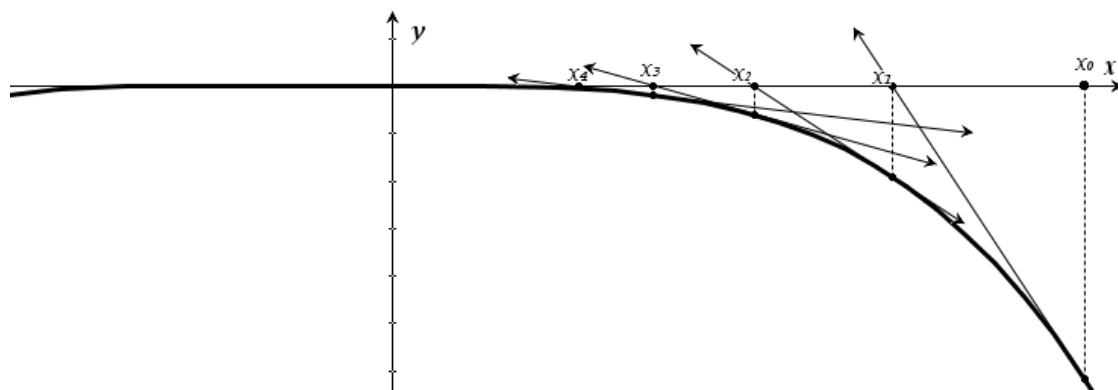


Fig. 3.41

n	x_n	$\frac{ e_{n+1} }{ e_n }$
0	1	0.72159023986073
1	0.72159023986073	0.72253049309681
2	0.52137095182042	0.71984890466253
3	0.37530830859079	0.71504809348557
4	0.26836349052714	0.70896981301523
5	0.19026161369914	0.70225676492774
6	0.13361250532629	0.69548345417247
7	0.092925286724966	0.68914790617628
8	0.064039266777341	0.68361279513245
9	0.043778062159891	0.67906284693452
10	0.029728055523572	0.67551285752372
11	0.020081683735352	0.67285828611075
12	0.013512127300387	0.67093770151532
13	0.009065795633504	0.66958192784852
14	0.0060702929177623	0.66864171920994
15	0.0040588510926405	0.66799782087484
16	0.0027113036851393	0.66756066567638
17	0.0018099596929024	0.66726562286717
18	0.001207723881849	0.66706730395089
19	8.0563311378212E-4	0.66693437505743
20	5.3730441726585E-4	0.66684539809824
21	3.5829897803159E-4	0.66678603917398
22	2.3890875640177E-4	0.66674713629467
23	1.592917291666E-4	0.66672160325935
24	1.0620323705591E-4	0.66669954731464
25	7.0805650068525E-5	

Tabela 3.18

Da análise da tabela percebe-se a necessidade de vinte e cinco iterações para obter uma aproximação correta com erro absoluto inferior a 0.5×10^{-4} e que a convergência não será quadrática, mas linear, com constante de erro assintótico $\frac{2}{3}$, ou seja, em cada iteração e_n decresce apenas 33% em relação ao erro anterior. O número de 25 iterações realizadas é fácil de entender uma vez que a partir de um erro inicial $e_0 = 1$ tem-se $e_n \approx \left(\frac{2}{3}\right)^n$ e da expressão

$$\left(\frac{2}{3}\right)^n < 0.5 \times 10^{-4}$$

resulta

$$n > \frac{\log_{10} 0.5 - 4}{\log_{10} \left(\frac{2}{3}\right)} \approx 24.4$$

A razão de ser tão lenta a convergência neste caso é a seguinte: se $f \in C^m[a, b]$ e r é um zero de f de multiplicidade m , então o método de Newton-Raphson converge linearmente desde que a estimativa inicial, em $[a, b]$, esteja “suficientemente” próxima de r [30, p. 41].

Alguns autores apresentam modificações do método de Newton-Raphson que permitem acelerar a convergência quando uma raiz tem multiplicidade $m > 1$. Em [16, p. 57] é apresentada a fórmula iteradora

$$x_{n+1} = x_n - \frac{mf(x_n)}{f'(x_n)}. \quad (40)$$

Usando esta fórmula no caso da equação do exemplo 3.9 obtemos os resultados que constam da tabela 3.19.

n	x_n
0	1
1	0.16477071958224
2	0.01620733771144
3	0.00024654143774
4	0.00000006072272
5	-0.0000000633250

Tabela. 3.19

A convergência quadrática é observável nas primeiras quatro iterações, em que o número de casas decimais corretas praticamente duplica em cada iteração, até ao valor de x_4 . Os valores de x_6 e x_7 são idênticos ao valor de x_5 . A razão pela qual não é possível obter maior precisão na aproximação é o mau condicionamento da raiz, isto é, ela é muito sensível a pequenos erros produzidos pelo cálculo da função. Este problema é do foro da teoria da perturbação e sobre isto nada mais diremos neste trabalho. O leitor mais interessado pode consultar, por exemplo, [16, p. 43].

Se aplicarmos a fórmula (40) à equação $(2x - 1)^5 = 0$ do exemplo 3.7, em que a raiz tem multiplicidade 5, obtemos a solução exata logo à 1ª iteração, independentemente da aproximação inicial:

$$x_{n+1} = x_n - \frac{5(2x_n - 1)^5}{10(2x_n - 1)^4} = x_n - \frac{1}{2}(2x_n - 1) = \frac{1}{2}$$

Na realidade, se a função f que consta de (36) for da forma $(ax + b)^m$, $a \neq 0$, a expressão (40) fica:

$$x_{n+1} = x_n - \frac{m(ax + b)^m}{am(ax + b)^{m-1}} = x_n - \frac{1}{a}(ax_n + b) = -\frac{b}{a}$$

Percebe-se assim porque esta modificação do método de Newton-Raphson tinha mesmo que levar à solução exata logo à 1ª iteração. Claro que tem mais interesse a utilização da fórmula (40) para funções que não são da forma simples $(ax + b)^m$, como se ilustra a seguir.

Exemplo 3.10

Para resolver a equação $(2x - 1)^5(x^2 - \sin x) = 0$ utilizando a fórmula (40), com $x_0 = 1$, pode desenvolver-se o processo iterativo na aplicação *Listas e Folha de Cálculo* da calculadora gráfica TI-Nspire. Obtemos assim as iterações que podemos observar na tabela 3.20, revelando a rapidez da convergência.

n	x_n
0	1
1	0.73968859120801
2	0.40689641106243
3	0.50079418513711
4	0.49999993157275
5	0.5

Tabela 3.20

Uma limitação importante da fórmula (40) é que, em geral, não se conhece a multiplicidade m da raiz. Esta dificuldade é ultrapassada na modificação que se descreve a seguir [22, p. 84]. Sendo r um zero de f com multiplicidade m , com

$$f(x) = (x - r)^m q(x)$$

será r um zero simples da função μ , definida por $\mu(x) = \frac{f(x)}{f'(x)}$.

A função iteradora (38)

$$g(x) = x - \frac{\mu(x)}{\mu'(x)} \quad (41)$$

dá, neste caso,

$$x_{n+1} = x_n - \frac{f(x_n)f'(x_n)}{[f'(x_n)]^2 - f(x_n)f''(x_n)}. \quad (42)$$

Tal como no caso da fórmula (40), este procedimento aplicado a uma função simples do tipo $(ax + b)^m$, $a \neq 0$, “lineariza” a função e produz a raiz exata $-\frac{b}{a}$.

$$\begin{aligned} x_{n+1} &= x_n - \frac{am(ax_n + b)^{2m-1}}{a^2m^2(ax_n + b)^{2m-2} - a^2m(m-1)(ax_n + b)^{2m-2}} = \\ &= x_n - \frac{am(ax_n + b)^{2m-1}}{a^2m(ax_n + b)^{2m-2}} = x_n - \frac{1}{a}(ax_n + b) = -\frac{b}{a} \end{aligned}$$

Mas, em geral, as expressões (40) e (42) não são equivalentes como se pode concluir da sua aplicação ao caso da equação $(2x - 1)^5(x^2 - \sin x) = 0$, usada no exemplo 3.10, cujos resultados se podem observar na tabela 3.21.

n	x_n com (26)	x_n com (40)	x_n com (42)
0	1	1	1
1	0.947937718242	0.73968859120801	0.77886253449177
2	0.909999142102	0.40689641106243	0.72544445291165
3	0.886983685236	0.50079418513711	0.60701218535109
4	0.878026285857	0.49999993157275	0.50818965142444
5	0.876750217079	0.5	0.50000962396293
6	0.876726223754	n.d.	0.50000000000987
7	0.876726215395	n.d.	0.5
8	0.876726215395	n.d.	n.d.

(n.d. – não definido)

Tabela 3.21

Com $x_0 = 1$, a fórmula (42) requer mais duas iterações que a fórmula (40) para chegar à solução 0.5. Também se pode observar que, com este valor inicial, o método de Newton-Raphson converge rapidamente para a raiz da equação $x^2 - \sin x = 0$ que está mais próxima de 1. Modificando o valor inicial para $x_0 = 0.45$, por exemplo, obtém-se convergência relativamente lenta para 0.5 com o método básico de Newton-Raphson e a habitual rapidez com os métodos modificados. Na tabela 3.22 podemos apreciar a evolução das iterações produzidas com as três fórmulas a partir de $x_0 = 0.95$ e de $x_0 = 0.88$.

n	x_n com (26)	x_n com (40)	x_n com (42)
0	0.95	0.95	0.95
1	0.91140428383214	0.75702141916071	0.82784566247622
2	0.88770077431986	0.3579529024252	0.83619217760305
3	0.87819909978816	0.50412429427242	0.84917872370921
4	0.87675693362839	0.49999802756918	0.86439721819695
5	0.87672622908452	0.49999999999958	0.87438971836963
6	0.87672621539507	0.5	0.87664604841773
7	0.87672621539506	n.d.	0.87672612207227
8	0.87672621539506	n.d.	0.87672621539493
0	0.88	0.88	0.88
1	0.8768735662879	0.86436783143951	0.87657324081456
2	0.87672652976286	0.94020069716482	0.87672587546119
3	0.87672621539649	0.76336832553254	0.87672621539339
4	0.87672621539506	0.33163883847874	0.87672621539506
5	0.87672621539506	0.50776303351899	0.87672621539506
6	0.87672621539506	0.4999925155663	0.87672621539506
7	0.87672621539506	0.49999999999402	0.87672621539506
8	0.87672621539506	0.5	0.87672621539506

(n.d. – não definido)

Tabela 3.22

É de notar que as iterações produzidas com a fórmula (40) convergem para a raiz 0.5 mesmo com $x_0 = 0.88$ que está muito próxima doutra raiz da equação. Mesmo com valores iniciais ainda mais próximos da raiz 0.87672621539506 observámos convergência para 0.5.

Como refere em [22, p. 84], a função g , definida em (41), tem as condições de continuidade necessárias para garantir convergência quadrática, independentemente da multiplicidade do zero de f . A única desvantagem deste método modificado parece ser o maior custo computacional, de que se destaca o cálculo da 2ª derivada; hipoteticamente também pode ocorrer cancelamento subtrativo no cálculo do denominador de (42).

Critério de convergência

Na prática, e como se viu, o teorema 3.5 não é prático pois nada diz sobre quão próximo tem de ser x_0 da raiz para garantir a convergência. Mas esta garantia existe sempre que f é duas vezes continuamente diferenciável num intervalo $[a, b]$ ao qual pertence x_0 e cumprem-se as condições que se apresentam a seguir [17, p. 42]:

1. $f(a) \times f(b) < 0$
(garantia da existência de solução)
2. $f'(x) \neq 0, x \in [a, b]$
(garantia da unicidade da solução)
3. $f''(x) \geq 0$ ou $f''(x) \leq 0$ em $[a, b]$
(f'' não muda de sinal em $[a, b]$, ou seja, a função é aí côncava ou convexa)
4. $\left| \frac{f(a)}{f'(a)} \right| < b - a$ e $\left| \frac{f(b)}{f'(b)} \right| < b - a$

Como ilustração do critério anterior, consideremos a equação $2^x - x^2 = 0$.

A função $f(x) = 2^x - x^2$ e as derivadas f' e f'' são contínuas e deriváveis, e as respetivas representações gráficas podem ser observadas nas figuras 3.42 e 3.43.

O gráfico de f sugere que a menor das raízes pertence a $[-1, -0.5]$ (as outras raízes são 2 e 4, como é fácil constatar).

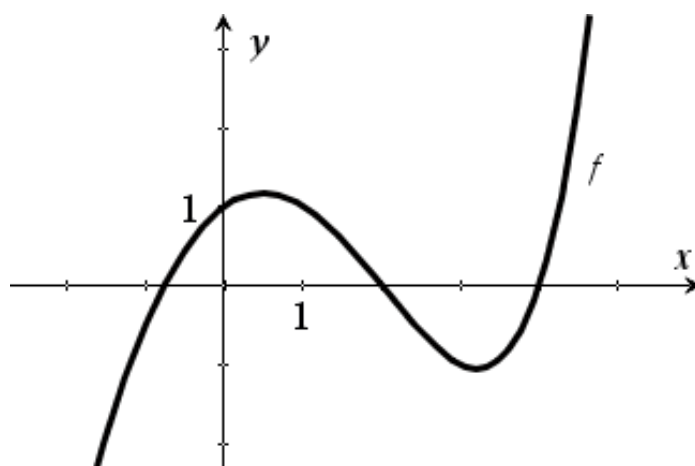


Fig. 3.42

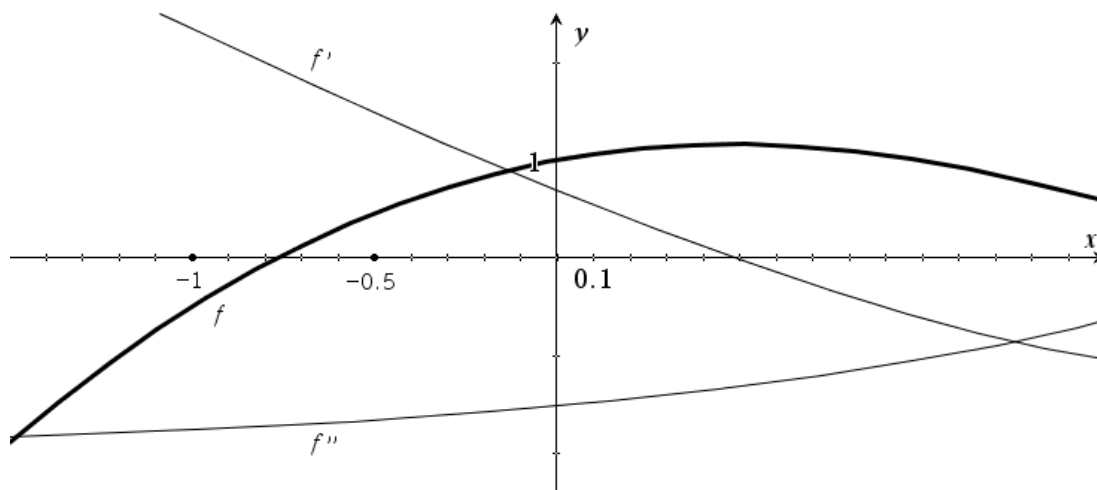


Fig. 3.43

Tem-se:

1. $f(-1) \times f(-0.5) = -0.5 \times (0.5\sqrt{2} - 0.25) \approx -0.5 \times 0.46 < 0$
2. $f'(x) = \ln 2 \times 2^x - 2x$ não se anula no intervalo considerado, como se pode perceber pelo comportamento gráfico de f'
3. $f''(x) = (\ln 2)^2 \times 2^x - 2 < 0$ no mesmo intervalo $[-1, -0.5]$
- 4.

$$\left| \frac{f(a)}{f'(a)} \right| = \left| \frac{f(-1)}{f'(-1)} \right| \approx 0.3 < 0.5 = |-0.5 - (-1)| = b - a$$

$$\left| \frac{f(b)}{f'(b)} \right| = \left| \frac{f(-0.5)}{f'(-0.5)} \right| \approx 0.3 < 0.5 = |-0.5 - (-1)| = b - a$$

Na tabela 3.23 apresentamos os resultados obtidos com os valores $x_0 = -0.99$ e $x_0 = -0.1$; no primeiro caso, a convergência é garantida pelo critério anterior com o intervalo $[-1, -0.5]$, por exemplo; no segundo caso, e uma vez que $x_0 = -0.1$ não pertence a $[-1, -0.5]$, a convergência não era certa (mas também neste caso é possível encontrar um intervalo que contem $x_0 = -0.1$ e no qual se cumprem as condições anteriores).

n	x_n	
	$x_0 = -0.99$	$x_0 = -0.1$
1	-0.78535188563175	-1.190115948776
2	-0.76681692380826	-0.82569622039066
3	-0.76666470621532	-0.76813639225333
4	-0.76666469596212	-0.76666565325292
5	-0.76666469596213	-0.76666469596252
6	-0.76666469596212	-0.76666469596212
7	-0.76666469596213	-0.76666469596213

Tabela 3.23

3.4.6. CONCLUSÃO

Do que se pode observar em toda a secção 3.4, comprova-se que este método é realmente muito eficiente. É relativamente simples em termos de desenvolvimento do algoritmo e, em geral, permite a rápida obtenção de boas aproximações da solução. Vimos também que um ou outro constrangimento, não todos, podem ser ultrapassados com algum cuidado na observação das condições de convergência ou com modificações do método.

A resolução de equações é problema que os métodos numéricos iterativos, não só os tratados neste trabalho, tentam dar resposta. No entanto, os diferentes comportamentos das funções na vizinhança dos seus zeros justificam o conhecimento e utilização de diferentes métodos. A compreensão das diferenças e ligações, como acontece com a expressão (38), que relaciona o método de Newton-Raphson com o método do ponto fixo, permite ajudar a tomar decisões sobre o método mais adequado para cada caso. Na verdade, não existe nenhum método que seja o melhor em todas as situações e é por esta razão que os melhores programas resultam de uma combinação de diferentes métodos.

Justifica-se assim neste trabalho uma análise comparativa, ainda que breve, que será feita na próxima secção.

3.5. COMPARAÇÃO ENTRE MÉTODOS NUMÉRICOS

3.5.1. INTRODUÇÃO

Na comparação dos três métodos desenvolvidos neste trabalho (bisseção, ponto fixo e Newton-Raphson) começamos por tecer algumas considerações sobre o método do ponto fixo.

Existe ampla liberdade na escolha da função iteradora φ , bastando para tal escrever a equação $f(x) = 0$ na forma matematicamente equivalente, num intervalo que contenha a raiz procurada $x = \varphi(x)$. A dificuldade maior, em geral, será a de selecionar φ (existem infinitas escolhas possíveis) que garanta convergência, tão rápida quanto possível, para a raiz desejada. De acordo com o teorema 3.3, quanto menor for o valor de $|\varphi'(r)|$ maior é a velocidade de convergência.

O método de Newton-Raphson pode ser entendido, neste contexto, como correspondendo à escolha particular da função iteradora $\varphi(x) = x - f(x)/f'(x)$, para a qual se tem $\varphi'(r) = 0$, que é o melhor possível.

Ambos os métodos têm convergência local, isto é, a convergência não é garantida a partir de uma aproximação inicial arbitrária x_0 e, em geral, depende de ser x_0 “suficientemente” próxima da raiz. Em ambos os casos existem critérios que podem ser usados na prática para garantia de convergência. A convergência do método de Newton-Raphson é quadrática enquanto que a convergência do método do ponto fixo é linear. Porém, há que fazer notar, a favor do método do ponto fixo, que ele não requer o cálculo de derivadas e esta é uma importante vantagem sobre o método de Newton-Raphson nos casos em que $f'(x)$ não exista ou tenha um custo computacional elevado.

O método da bisseção é, de todos os métodos de cálculo de raízes de equações, o mais robusto. De facto, a partir de um intervalo em que f seja contínua e tenha sinais diferentes nos extremos desse intervalo, a convergência para uma raiz nesse intervalo é garantida. Por outro lado, a convergência é lenta uma vez que em cada iteração se reduz para metade a amplitude do intervalo que contém a raiz. Tal não significa, porém, que não existam situações em que métodos de convergência geralmente mais rápida, como é o caso do método de Newton-Raphson, acabem por ser vencidos pelo método da bisseção mesmo em termos da velocidade de convergência, como ilustraremos à frente.

3.5.2. ALGUNS EXEMPLOS

Para simplificação de escrita, usaremos a notação seguinte para a sucessão das iterações produzidas com os distintos métodos: $(x_n)_b$ para o ponto médio dos sucessivos intervalos produzidos no método da bisseção, $(x_n)_{pf}$ para o método do ponto fixo e $(x_n)_{nr}$ para o método de Newton-Raphson. Nos gráficos usaremos, em alternativa, e com o mesmo significado, xn_b , xn_{pf} e xn_{nr} , respetivamente.

Exemplo 3.11

Como vimos no exemplo 3.5 deste capítulo, o método de Newton-Raphson não pode ser usado no caso da equação

$$\sqrt[3]{x} = 0.$$

Vamos usar o método do ponto fixo com a função iteradora

$$\varphi(x) = \frac{x}{1 - \sqrt[3]{x}}$$

cuja derivada é

$$\varphi'(x) = \frac{1 - \sqrt[3]{x} + \frac{x}{3\sqrt[3]{x^2}}}{(1 - \sqrt[3]{x})^2},$$

e pode concluir-se que $|\varphi'(x)| > 1$ para $x > 0$ e $|\varphi'(x)| < 1$ para $x < 0$, como ilustra a figura 3.44, na qual se representam graficamente estas funções φ e φ' .

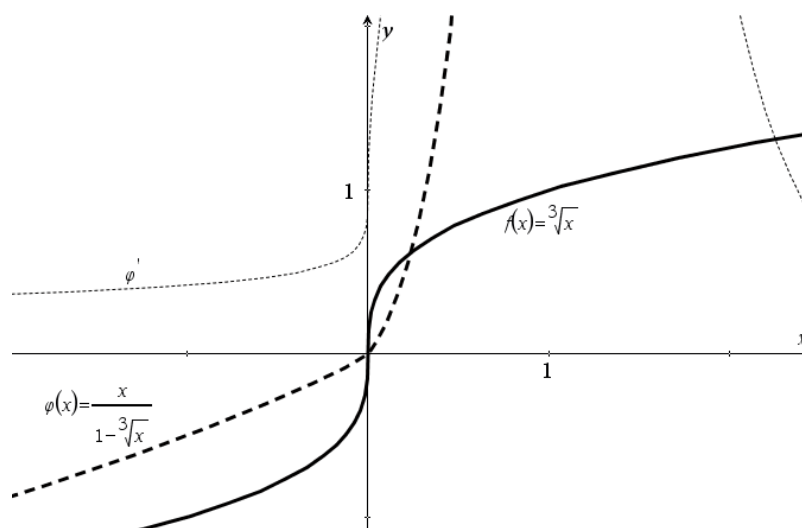


Fig. 3.44

Na tabela 3.24 apresentam-se as iterações produzidas com o método da bissecção, a partir do intervalo inicial $[-0.1, 0.01]$ e com a função iteradora escolhida, com a aproximação inicial $x_0 = -0.01$.

n	$(x_n)_b$	$(x_n)_{pf}$	n	$(x_n)_b$	$(x_n)_{pf}$
0		-0.01	38	1.091393639798E-13	-0.017557233760973
1	-0.045	0.0011111111111111	39	-9.0949470410275E-14	-0.013935336298593
2	-0.0175	0.0012394906303605	40	9.0949467847625E-15	-0.01123235367228
3	-0.00375	0.0013886589818596	41	-4.0927261812757E-14	-0.0091771117755865
4	0.003125	0.001563041243962
5	-3.125E-4	0.0017682524544741	60	8.4406487885725E-21	-7.8535534894499E-4
6	0.00140625	0.0020114912882646	61	-3.9264246800807E-20	-7.1901747982196E-4
7	5.46875E-4	0.0023020910202837	62	-1.5411799006118E-20	-6.5989891316693E-4
8	1.171875E-4	0.0026523026613747
9	-9.765625E-5	0.0030784272425183	99	-1.2196452091281E-31	-7.6267348930959E-5
10	9.765625E-6	0.0036024840467658	100	-3.5189821338435E-32	-7.3164595992142E-5
...
23	-3.5762786865238E-9	0.18906976609028	99	-1.1211142662017E-61	-5.5085407728951E-6
24	2.9802322387694E-9	0.44377361286844	200	-4.3658258563585E-62	-5.4129434610472E-6
25	-2.980232238772E-10	1.8705578458865
26	1.3411045074461E-9	-8.058183152282	499	-1.9604850825075E-152	-2.6197232233875E-7
27	5.2154064178445E-10	-2.6817372956891	500	1.3999449173555E-152	-2.6030672306932E-7

Tabela 3.24

Aparentemente qualquer destes métodos parece funcionar, ao contrário do método de Newton-Raphson, com vantagem clara do método da bissecção no que diz respeito à velocidade de convergência (na figura 3.45 faz-se uma representação gráfica das iterações calculadas com ambos os métodos). Observamos que enquanto os valores de $(x_n)_{pf}$ são positivos não ocorre convergência e tal acontece por ser, para estes valores, $|\varphi'(x)| > 1$. Tendo-se produzido uma aproximação negativa à 26ª iteração, o método convergiu muito lentamente para a raiz $r = 0$ por ser $|\varphi'(x)| < 1$. À medida que nos aproximamos de $r = 0$, a convergência é cada vez mais lenta já que está $|\varphi'(x)|$ a convergir para 1. De facto, comprovámos que ao fim de 1000 iterações se obtém $-2.4964423883153 \times 10^{-8}$.

Este exemplo ilustra bem as dificuldades que métodos geralmente mais rápidos que o da bissecção podem ter em certos casos e a superioridade do método da bissecção nestes casos, mesmo do ponto de vista da velocidade de convergência. A função iteradora e a aproximação inicial foram escolhidas justamente para ilustrar tais dificuldades e outras escolhas poderiam ter sido feitas com vantagem para o comportamento do método do ponto fixo.

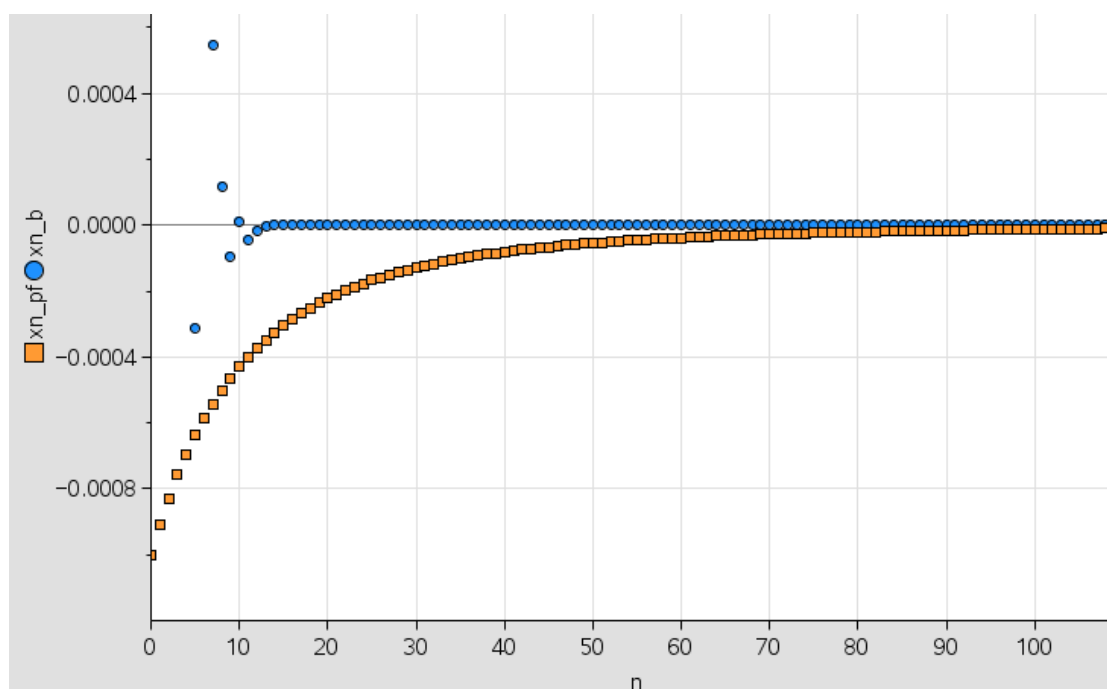


Fig. 3.45

Exemplo 3.12

Retomamos aqui a equação

$$\cos x = 0$$

que já usámos no exemplo 3.6. Vamos agora alargar a análise deste exemplo, comparando o desempenho dos três métodos, utilizando para o método do ponto fixo a função iteradora

$$\varphi(x) = x + 1 - 2 \sin^2 \frac{x}{2}.$$

Para o método da bissecção vamos partir do intervalo $[1,2]$, que contém apenas a menor solução positiva.

Na figura 3.46 pode observar-se o gráfico das funções que são necessárias para análise dos métodos e na tabela 3.25 apresentam-se alguns resultados. Com $x_0 = 0.364$, em apenas quatro iterações, a função φ escolhida produziu uma solução tão exata quanto a máquina permite para a raiz $r = \frac{\pi}{2}$ (uma vez que $x_5 = x_4$). A excelente convergência do método do ponto fixo neste caso é explicada por ser $\varphi'(x) = 1 - \sin x$ e $\varphi'\left(\frac{\pi}{2}\right) = 0$. Os quocientes $|e_{n+1}|/e_n^2$ tendem para uma certa constante $c \approx 0.64$ o que, de acordo com a definição 3.1 mostra que a

convergência é quadrática. O método de Newton-Raphson teve um início mais irregular e acabou por convergir para outra solução que não a mais próxima do valor de partida. É de enfatizar que os quocientes $|e_{n+1}|/e_n^2$ tendem para uma constante de erro assintótico diferente da anterior.

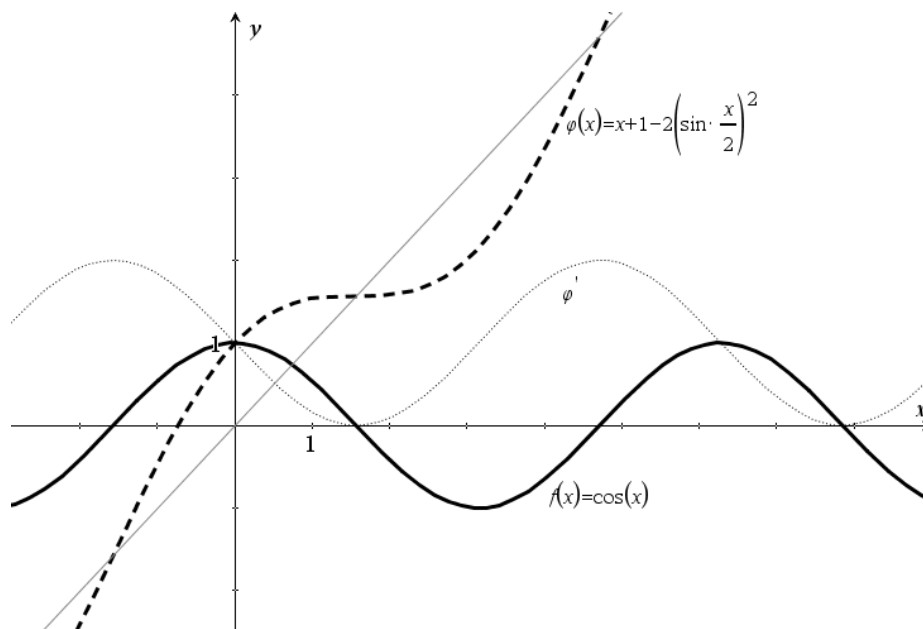


Fig. 3.46

n	Método da Bisseção	Método do Ponto Fixo		Método de Newton-Raphson	
	$(x_n)_b$	$(x_n)_{pf}$	$\frac{ e_{n+1} }{e_n^2}$	$(x_n)_{nr}$	$\frac{ e_{n+1} }{e_n^2}$
0		0.364	9.8001...	0.364	22.557...
1	1.5	1.2984802433378	0.92965...	2.9888339628807	-0.39252...
2	1.75	1.5674431364547	0.63934...	-3.5064394769262	-0.49813...
3	1.625	1.5707963205111	0.63661...	-6.1246070288753	0.0034251...
4	1.5625	1.5707963267949	0.63661...	0.1284786225665	476.71...
5	1.59375	1.5707963267949	0.63934...	7.8690012363013	0.12683...
6	1.578125			7.8539805044563	0.12732...
7	1.5703125			7.8539816339745	0.12732...
8	1.57421875			7.8539816339745	0.12732...
16	1.570785522461				
17	1.5707931518555				
18	1.5707969665528				
19	1.5707950592042				
20	1.5707960128785				

Tabela 3.25

Sem surpresa, o método da bisseção não consegue melhor do que 6 algarismos corretos ao cabo de vinte iterações.

Exemplo 3.13

Voltamos a uma situação em que o método de Newton-Raphson (sem qualquer modificação) não converge rapidamente. A equação

$$(2x - 1)^5 = 0$$

tem a solução única $r = 0.5$ cuja multiplicidade é cinco. Como vimos no exemplo 3.7, o método de Newton não consegue mais do que um algarismo significativo correto no final de dez iterações. Não é difícil perceber que o “lento” método da bisseção consegue melhor do que isto. Mas vale a pena confirmar esta ideia com um teste numérico e, para tal, usaremos o intervalo inicial $[0.1, 0.7]$. Na tabela 3.26 e nas figuras 3.47 e 3.48 pode observar-se o comportamento das iterações de ambos os métodos.

n	$(x_n)_b$	$(x_n)_{nr}$
0		0.2
1	0.55	0.26
2	0.325	0.308
3	0.4375	0.3464
4	0.49375	0.37712
5	0.521875	0.401696
6	0.5078125	0.4213568
7	0.50078125	0.43708544
8	0.497265625	0.449668352
9	0.4990234375	0.4597346816
10	0.49990234375	0.46778774528
26	0.4999998956922	0.49909330563529
27	0.4999999627474	0.49927464450823
28	0.49999996275	0.49941971560658
29	0.500000013039	0.49953577248526
30	0.500000004657	0.49962861798821

Tabela 3.26

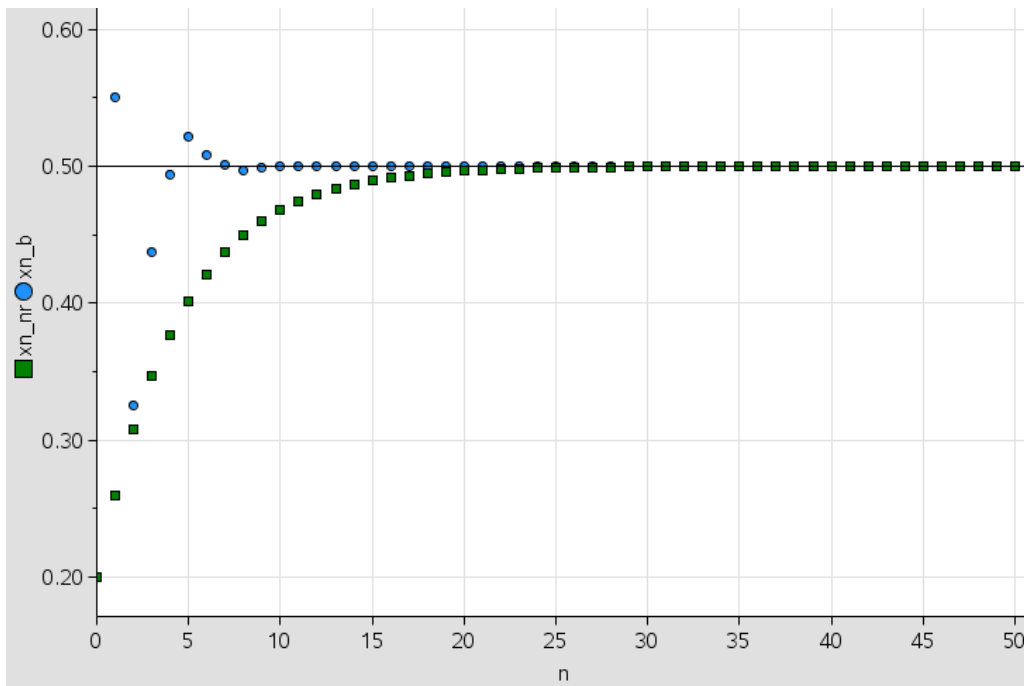


Fig. 3.47

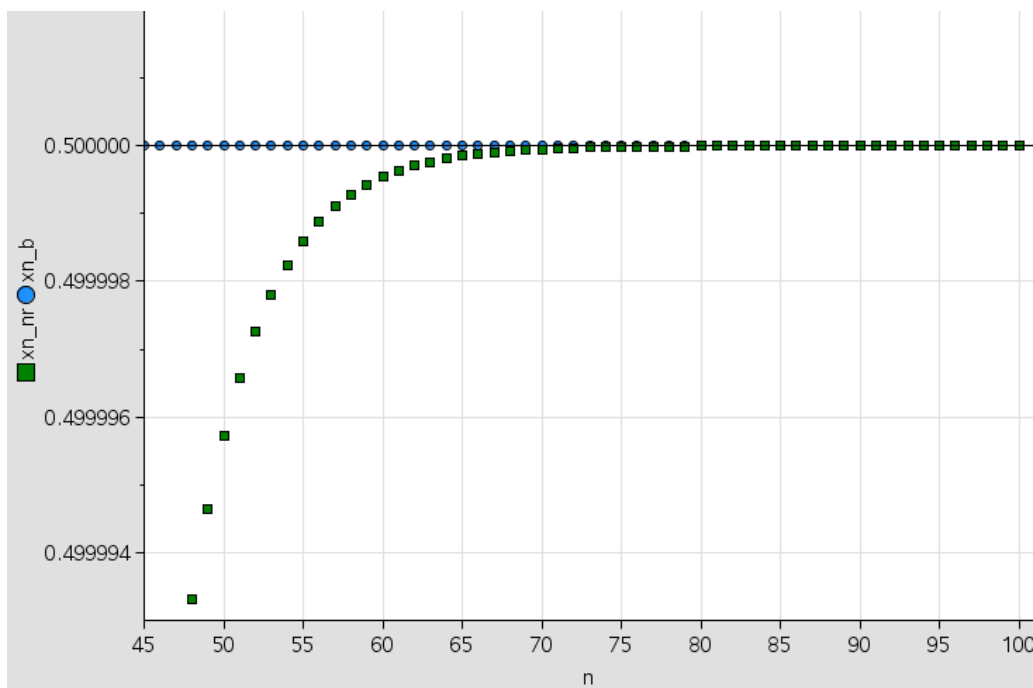


Fig. 3.48

A convergência do método de Newton-Raphson é tão mais lenta quanto maior é a multiplicidade da raiz. Com $f(x) = (x - a)^m$, tem-se

$$x_{n+1} = x_n - \frac{(x_n - a)^m}{m(x_n - a)^{m-1}} = x_n - \frac{x_n - a}{m}$$

e

$$e_{n+1} = e_n \left(1 - \frac{1}{m}\right).$$

Por aqui se pode ver que para $m = 2$ (raiz dupla), a convergência do método de Newton-Raphson é tão rápida quanto a do método da bisseção. Para valores superiores de m , o método da bisseção é mais rápido.

Exemplo 3.14

Finalmente, comparemos o comportamento dos métodos da bisseção e de Newton-Raphson no caso da equação

$$\arctg(x) = 0.$$

A figura 3.49 contém uma representação gráfica da função $f(x) = \arctg(x)$ e das respectivas derivadas, que são funções contínuas; além disto, é importante notar que a derivada

$$f'(x) = \frac{1}{1+x^2}$$

nunca se anula e toma valores muito próximos de um na vizinhança da raiz $r = 0$.

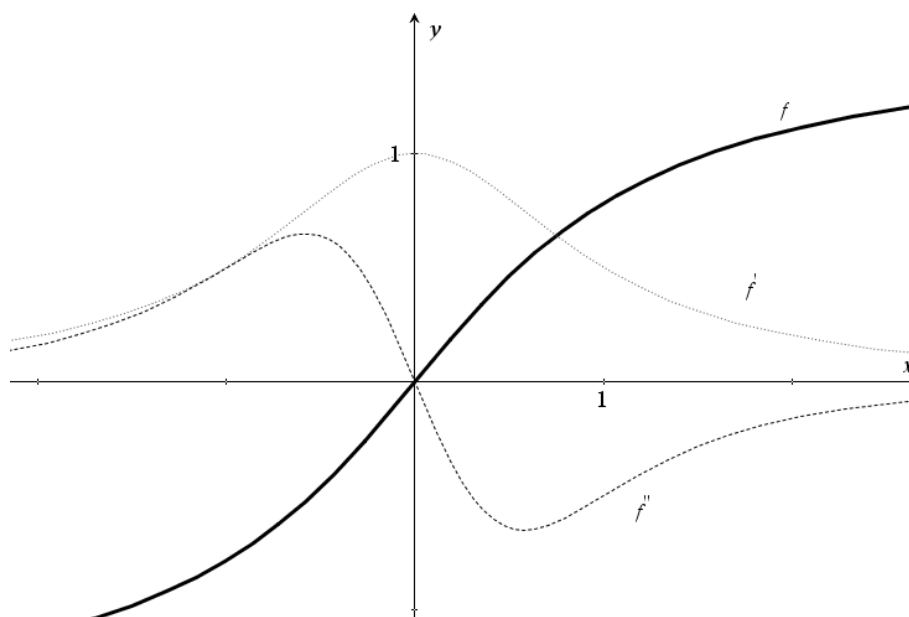


Fig. 3.49

Vamos considerar o intervalo $[-1,2]$ para iniciar a aplicação do método da bisseção e diferentes estimativas iniciais neste intervalo para aplicar o método de Newton-Raphson. Apresentam-se alguns resultados na tabela 3.27.

n	$(x_n)_b$	$(x_n)_{nr}$		
		0.2	1.38	1.4
0				
1	0.5	-0.00529138224388	-1.3609568191195	-1.4136186488037
2	-0.25	9.87674218E-8	1.311869363788	1.4501293146283
3	0.125	0.	-1.1900583669442	-1.5506259756376
4	-0.0625	0.	0.9168140525076	1.8470540841498
5	0.03125	0.	-0.4489229094892	-2.8935623931412
6	-0.015625	0.	0.05807286354739	8.7103258469748
7	0.0078125	0.	-1.30477570447E-4	-103.24977377171
8	-0.00390625	0.	1.48087E-12	16540.563827174
9	0.001953125	0.	0.	-429721482.89304
10	-9.765625E-4	0.	0.	2.9006411727663E17
26	-1.4901161193841E-8	0.	0.	n.d.
27	7.450580596931E-9	0.	0.	n.d.
28	-3.725290298455E-9	0.	0.	n.d.
29	1.862645149238E-9	0.	0.	n.d.
30	-9.313225746085E-10	0.	0.	n.d.

Tabela 3.27

Como é expectável, o comportamento do método de Newton-Raphson depende da aproximação inicial. Com $x_0 = 0.2$, a velocidade de convergência é notável, com $x_0 = 1.38$ observa-se um comportamento oscilatório das iterações com convergência quadrática a partir da sexta iteração. Mas com $x_0 = 1.4$ o método diverge. Não há razão para dizer que este é um comportamento anómalo do método de Newton-Raphson porque $x_0 = 1.4$ não está “suficientemente” próximo da raiz.

3.5.3. CONCLUSÃO

Como vimos, o mais robusto dos métodos estudados é o método da bissecção. Desde que se parta de um intervalo onde a função seja contínua e em cujos extremos tome sinais opostos, a convergência para uma raiz é garantida. Não admira assim que o método tenha calculado uma raiz em todos os exemplos em que o testámos. No entanto, o método da bissecção é, em geral, mais lento que os outros métodos. Também é de notar que nem sempre a raiz está num intervalo em cujos extremos a função toma sinais opostos (um exemplo simples é a equação $x^2 = 0$).

O método do ponto fixo é um método cuja eficácia depende de uma boa escolha da função iteradora φ . De facto, para obter convergência para uma certa raiz r , é necessário que a raiz esteja contida num intervalo onde se cumpra a condição $|\varphi'(x)| < 1$. Embora, como vimos no exemplo 3.12, seja possível obter convergência quadrática, esta é normalmente apenas linear e o método de Newton-Raphson será, em geral, mais rápido.

3.6. O CÁLCULO DE RAÍZES EM MÁQUINAS DE CALCULAR

As máquinas de calcular com capacidades gráficas, como as três que constam deste trabalho, implementam métodos numéricos para resolver equações. Porém, os manuais de utilização destas máquinas não são elucidativos relativamente aos métodos implementados, os códigos que implementam esses métodos não são “abertos” e os nossos pedidos de informação aos fabricantes Casio e a Texas Instruments também se revelaram infrutíferos. Desta maneira, no contexto deste trabalho, tudo o que podemos fazer é comparar a performance de cada uma das calculadoras na determinação de raízes de algumas equações e fazer conjecturas sobre os métodos que usam.

Vejamos como se comportam as calculadoras TI84 Plus C SE, fx-CG20 e TI-Nspire Cx para “resolver” a equação $x^2 = 0$.

TI84 Plus C SE

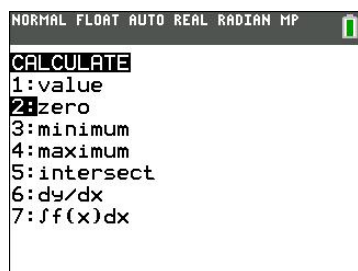


Fig. 3.62

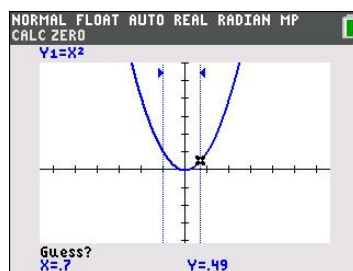


Fig. 3.63

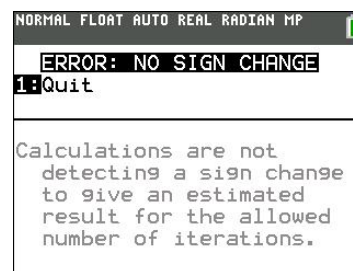


Fig. 3.64

Na máquina TI84 (ver figuras 3.62, 3.63 e 3.64), a funcionalidade “zero” não consegue calcular a raiz da equação $x^2 = 0$. Como se percebe da figura 3.64, tal acontece por não ser possível encontrar um intervalo em cujos extremos f tenha sinais contrários. Isto não significa que a calculadora TI-84 Plus C SE use o método da bisseção tal como o apresentamos (e de exemplos que daremos mais adiante podemos concluir que não é o método da bisseção que a TI-84 usa) As calculadoras Casio fx-CG20 e TI-Nspire Cx são mais eficazes, que a TI84 na resolução desta equação, como se pode observar nas figuras 3.65 a 3.68.

Casio fx-CG20

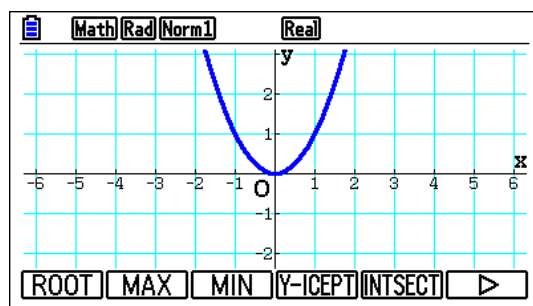


Fig. 3.65

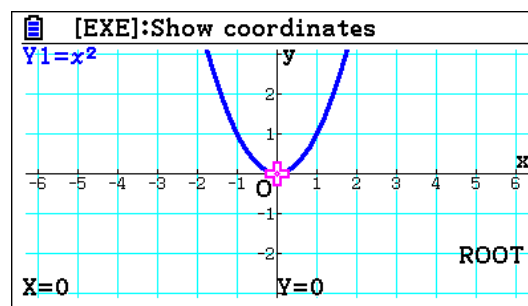


Fig. 3.66

TI-Nspire CX

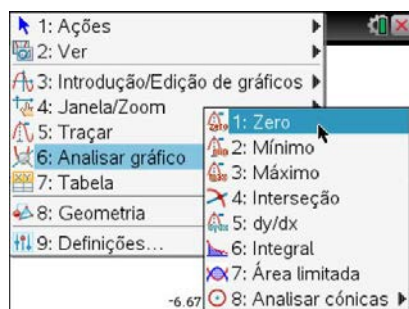


Fig. 3.67

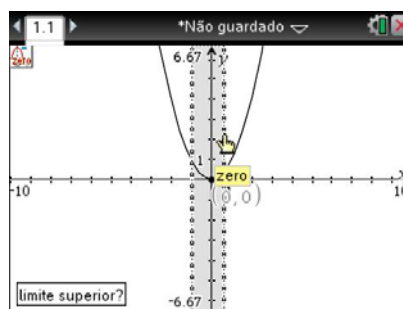


Fig. 3.68

No modelo Casio não é necessário, ao contrário dos modelos TI, definir valores iniciais à esquerda e à direita da solução. Nos modelos TI, sendo necessário definir um intervalo inicial contendo a solução, não parece que seja implementado o método da bisseção, pois a solução é encontrada.

Na resolução numérica de uma equação $f(x) = 0$, a continuidade de f é condição indispensável. Os exemplos que se seguem são bem elucidativos desta necessidade.

Na tabela 3.33 apresentam-se as primeiras 20 aproximações obtidas pelo método da bisseção a partir do intervalo $[1,2]$ no caso da equação

$$\tan(x) = 0 \quad (48)$$

Como se pode apreciar, a sequência está a convergir para $\frac{\pi}{2}$ que não é uma raiz da equação, mas sim um ponto de descontinuidade da função tangente (ver figura 3.69).

i	p_i
1	1,50000000000000
2	1,75000000000000
3	1,62500000000000
4	1,56250000000000
5	1,59375000000000
6	1,57812500000000
7	1,57031250000000
8	1,57421875000000
9	1,57226562500000
10	1,57128906250000

i	p_i
11	1,57080078125000
12	1,57055664062500
13	1,57067871093750
14	1,57073974609375
15	1,57077026367187
16	1,57078552246094
17	1,57079315185547
18	1,57079696655273
19	1,57079505920413
20	1,57079601287844

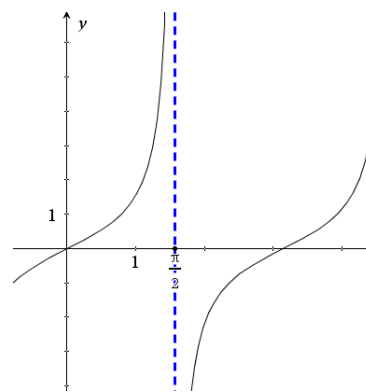


Tabela 3.33

Fig. 3.69

Na realidade, partindo do intervalo $[1,2]$, cada passo de bissecção produz um novo intervalo contendo $\frac{\pi}{2}$ uma vez que a função tangente é positiva à esquerda de $\frac{\pi}{2}$ e negativa à direita desse ponto. Para terminar este exemplo, observemos que as raízes da equação $\tan(x) = 0$ são afinal as raízes da equação $\sin(x) = 0$. Uma vez que a função seno é contínua em todo o seu domínio, ao resolver esta última equação não ocorrerá confusão entre raízes e pontos de descontinuidade.

No Matlab, a função `fzero` é “enganada” por esta descontinuidade. Com efeito

```
>> fzero(@tan,[1,2]) produz o resultado ans = 1.5708.
```

Como se pode observar nas figuras 3.70 a 3.75, as calculadoras gráficas estão mais “atentas” ao problema da continuidade.

T184 Plus C SE

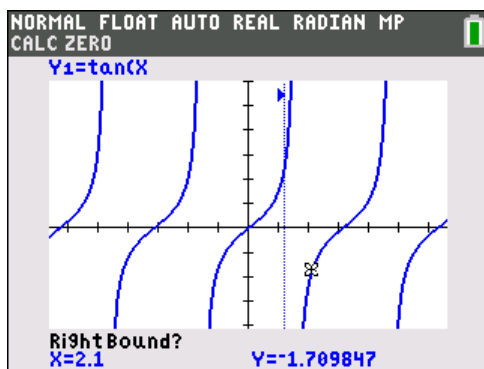


Fig. 3.70

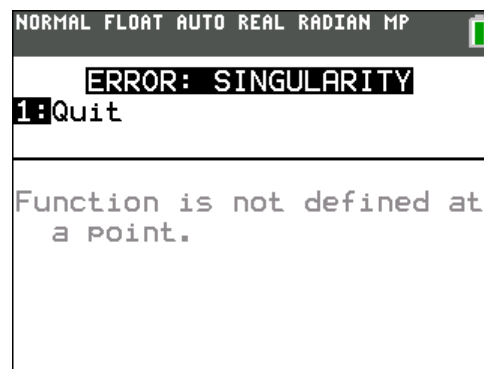


Fig. 3.71

Casio fx-CG20

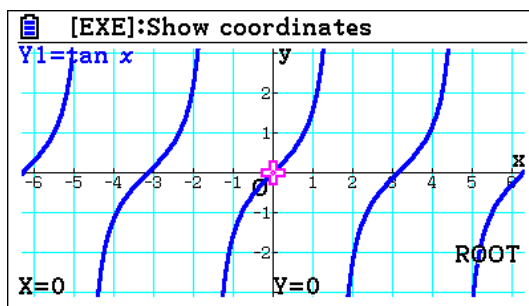


Fig. 3.72

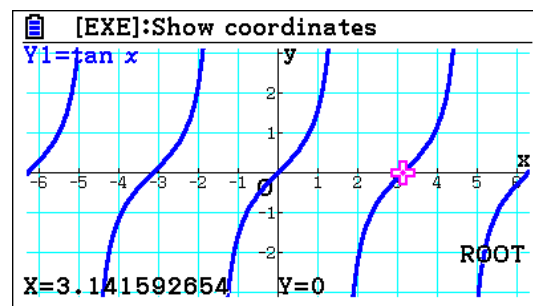


Fig. 3.73

TI-Nspire CX

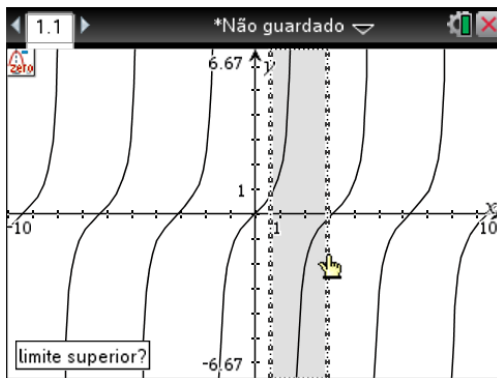


Fig. 3.74

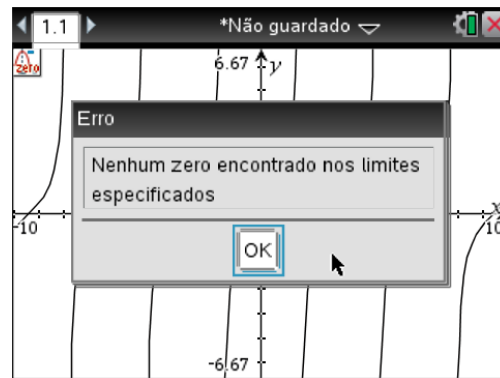


Fig. 3.75

No modelo Casio a funcionalidade percorre as soluções, da menor para a maior de entre as visíveis no ecrã, não se detendo em $\frac{\pi}{2}$. Nos modelos da TI, como já se disse, é necessária a indicação do intervalo inicial e esta é a razão de não ser encontrada nenhuma raiz (que não existe no intervalo fixado).

Como segundo exemplo, consideremos a equação

$$x \sin\left(\frac{1}{x}\right) \quad (49)$$

E vejamos como se comportam as mesmas calculadoras.

TI84 Plus C SE

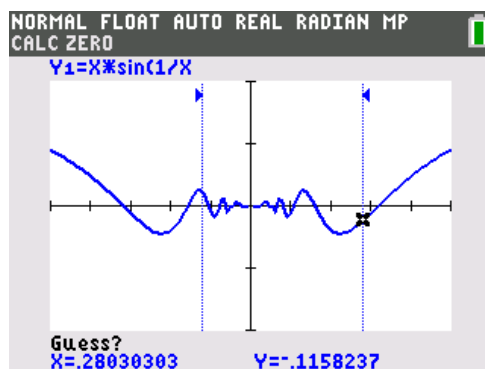


Fig. 3.76

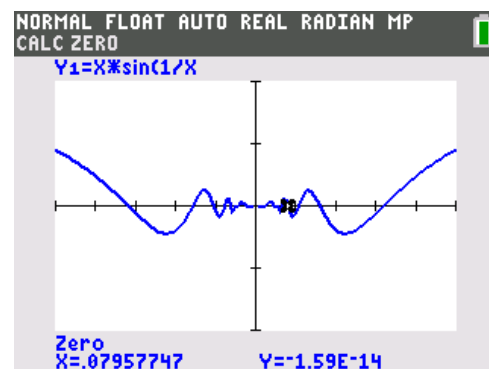


Fig. 3.77

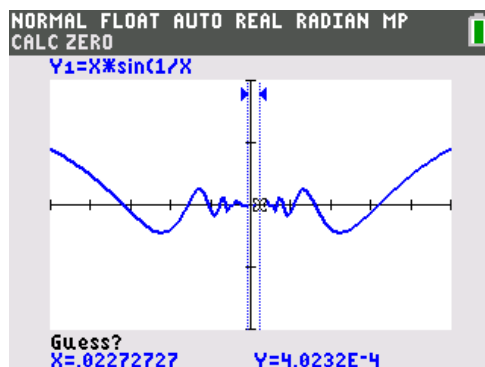


Fig. 3.78

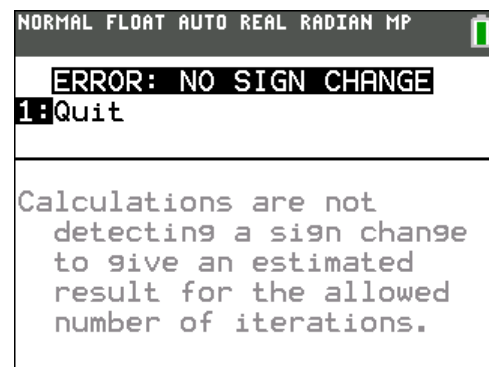


Fig. 3.79

Note-se que há uma infinidade de soluções em qualquer intervalo que contenha 0, da forma $\frac{1}{k\pi}$, $k \in \mathbb{Z} \setminus \{0\}$. Dependendo do intervalo inicial, é encontrada uma raiz (ver figura 3.77, onde se usou o intervalo $[-0.121, 0.280]$ e se encontra uma aproximação de $\frac{1}{4\pi}$) ou não, como acontece no caso do intervalo inicial $[-0.00758, 0.0227]$ a que diz respeito a figura 3.79.

Na janela de visualização utilizada anteriormente, $[-0.5, 0.5] \times [-1, 1]$, é possível ajustar ligeiramente o último intervalo inicial considerado e além do problema deixar de ocorrer, consegue-se um valor aproximado da solução $\frac{1}{80\pi}$, como se pode observar na figura 3.80.

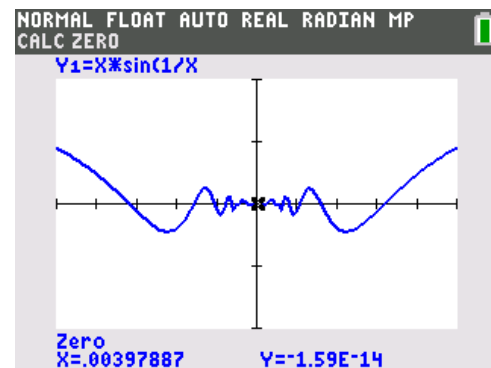


Fig. 3.80

Analisando mais exemplos, mesmo com janelas de visualização com diferentes escalas, observa-se uma regularidade que parece justificar esta diferença de comportamento, que não ocorre com outras funções nesta mesma máquina de calcular. Nos intervalos em que os extremos têm o mesmo sinal, é devolvida uma mensagem como a da figura 3.79 quando se pretende obter um zero da função no seu interior, mesmo que esses intervalos não contenham o ponto de descontinuidade. Nas figuras 3.81 e 3.82 pode observar-se que considerando um intervalo inicial em que a função tem o mesmo sinal nos seus extremos, $[-0.15, 3.12]$, é possível obter soluções, reforçando a diferença de comportamento do algoritmo utilizado pela máquina, pelo menos para esta função (49).

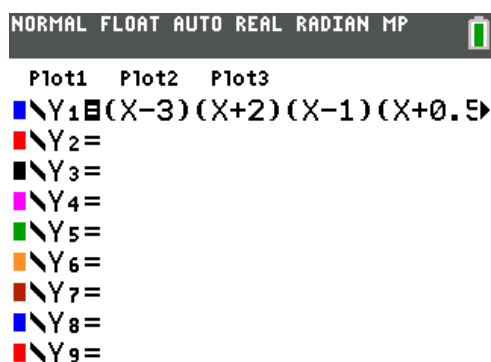


Fig. 3.81

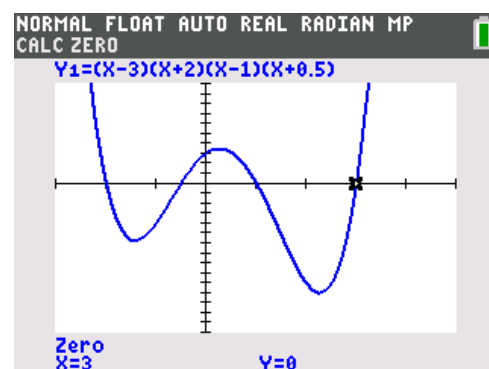


Fig. 3.82

Na Casio fx-CG20, que determina automaticamente as raízes sem que o utilizador defina previamente um intervalo inicial, encontra aproximações de todas as raízes para $|k| \leq 17$ e para valores de $|k|$ até 241 encontra ainda 12 raízes, para valores de $|k|$ cada vez mais distantes (20,

22, 24, 27, 30,34,40,48,60,80,120 e 241), sendo as raízes de menor valor absoluto encontradas as que se aproximam de $-\frac{1}{241\pi}$ e $\frac{1}{241\pi}$, numa vizinhança de 0 e com janela de visualização $[-0.5,0.5] \times [-1,1]$.

Casio fx-CG20

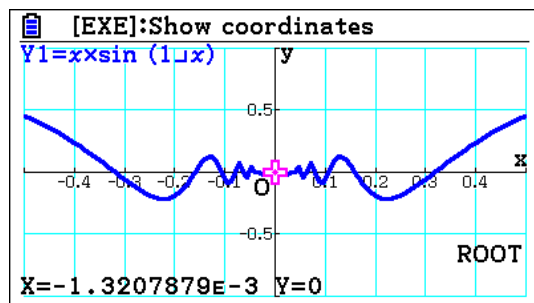


Fig. 3.83

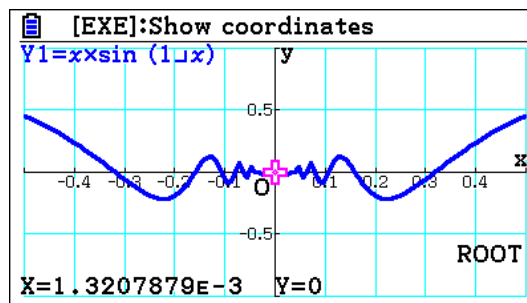


Fig. 3.84

O espaçamento entre valores sucessivos das abcissas é determinado pela máquina, em função da janela de visualização, sem que se conheça essa informação, mas é plausível que conduza a um erro de discretização que não permita capturar mais nenhuma das raízes (em número infinito) que existem entre as encontradas. Se definirmos uma nova janela de visualização este erro de discretização poderá ser menor pelo facto do espaçamento entre os valores sucessivos das raízes o poder também ser, o que poderá explicar a obtenção de soluções de menor valor absoluto, como se pode observar nas figuras 3.85 e 3.86, em que se considera uma janela de visualização diferente, $[-0.016,0.016] \times [-0.05,0.05]$, que contém as raízes em que $|k| \geq 20$. Observa-se que a máquina encontra muito mais raízes, algumas que não encontrou na janela anterior e não encontrou outras que tinha encontrado, sendo as mais pequenas em módulo as aproximadas a $-\frac{1}{7520\pi}$ e a $\frac{1}{7520\pi}$, num padrão análogo ao anterior, em que à medida que se aproximam de 0 são cada vez mais espaçadas. Confirma-se assim que a janela de visualização é determinante na obtenção das raízes. Nas figuras 3.85 e 3.86 podem observar-se as raízes mais próximas de 0 que esta máquina encontra nesta janela de visualização.

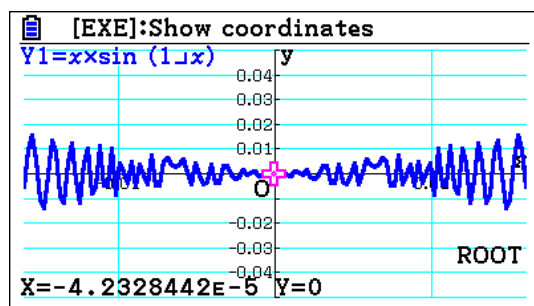


Fig. 3.85

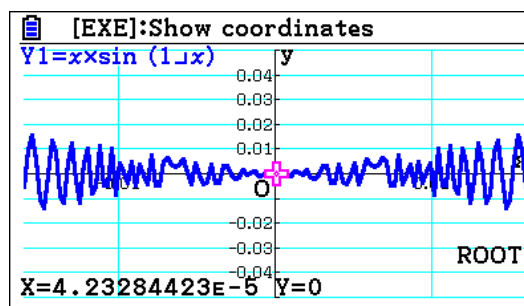


Fig. 3.860

No modelo TI-Nspire, aparentemente é devolvido sempre o mesmo valor para solução da equação (49), uma solução que é uma aproximação da solução exata na forma $\frac{1}{k\pi}$, com $k > 11,65 \times 10^6$. A grande variedade de testes efetuados, dos quais se observam 3 exemplares nas figuras 3.87, 3.88 e 3.89, leva a admitir que esta solução não depende, no geral, da janela de visualização nem do intervalo inicial que se considere, desde que a contenha. Note-se que na figura 3.89, em que o zero se observa graficamente distante da origem, o ponto que lhe corresponde não se encontra no eixo das abscissas. Caso o intervalo inicial não contenha esta raiz, observam-se outras, que variam de acordo com o intervalo inicial considerado, sendo muitas vezes 0 e outras vezes outros valores, como se pode observar no exemplo da figura 3.90.

TI-Nspire CX

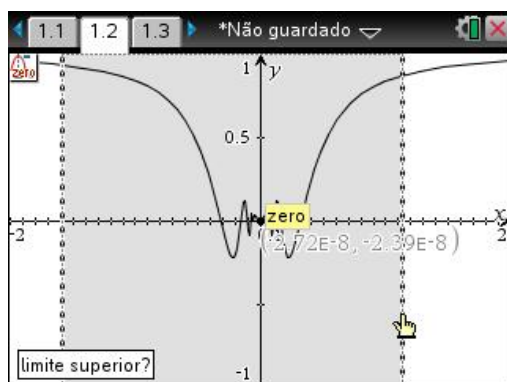


Fig. 3.87

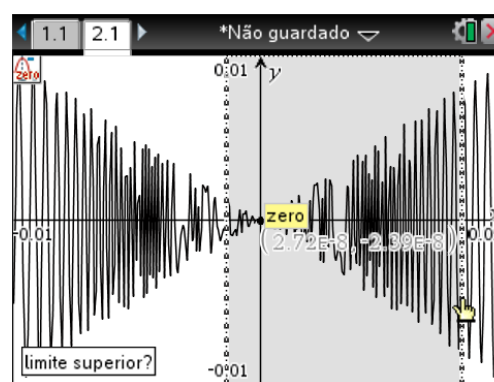


Fig. 3.88

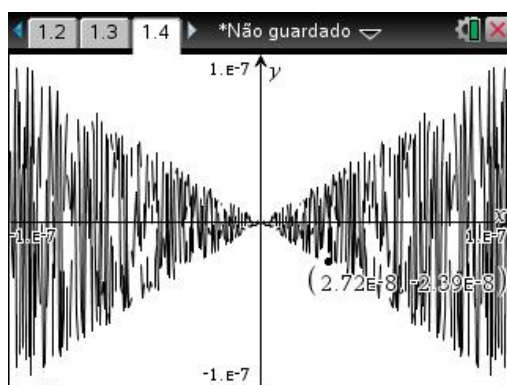


Fig. 3.89

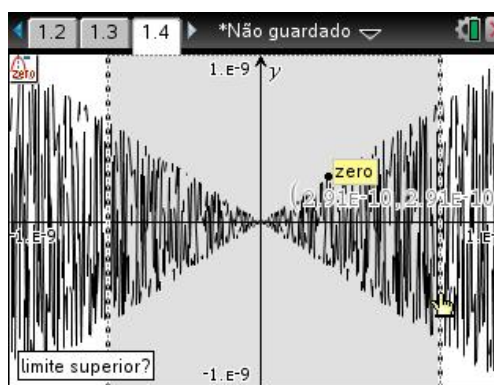


Fig. 3.90

Os mesmos testes revelaram que a dificuldade surgida na TI84 Plus C SE para esta função (49) quando a função toma o mesmo sinal nos extremos do intervalo inicial não ocorre com este modelo.

4. CONCLUSÕES E PERSPETIVAS DE INTEGRAÇÃO CURRICULAR

4.1. INTRODUÇÃO

Nesta secção pretendemos analisar a possibilidade de integração curricular nos programas do Ensino Secundário dos métodos numéricos apresentados antes. Faremos isto sobretudo através da apresentação de sugestões concretas de tarefas para os alunos. O foco será o tema matemático do programa oficial e a forma como aí se podem enquadrar, com vantagem, os métodos numéricos que propomos. Não é nosso objetivo, apesar da relevância, tratar de questões do foro da metodologia do ensino e da didática. Tal não significa que não se façam algumas considerações e que não fique expressa uma tendência em relação às orientações didáticas e metodológicas. Sabe-se que em relação a estas questões há opiniões divergentes com abrangência significativa, mas não será por isso que ficará diminuída a nossa proposta.

4.2. ALGUMAS CONSIDERAÇÕES GERAIS

Em Portugal, as experiências de integração de métodos numéricos no Ensino Secundário são de abrangência limitada. Uma experiência à oportunidade da introdução de métodos numéricos no Ensino Secundário foi vivenciada pelo autor com os seus alunos do 12º ano, como já foi referido, a propósito da resolução da equação $2^x = x^2$ [1]. Esta equação foi apresentada pelo coordenador dos Programas de Matemática do Ensino Secundário e do grupo de Acompanhamento, Professor Jaime Carvalho e Silva, no âmbito do estudo comparativo dos crescimentos da função quadrática e da função exponencial utilizando a calculadora gráfica (para uma apreciação deste estudo veja-se [31]). As vantagens destas abordagens elementares de problemas do foro da análise numérica a um nível pré-universitário são reconhecidas por muitos educadores, talvez não tanto em Portugal como noutros países, em particular os Estados Unidos da América e a Austrália.

Alguma da pesquisa bibliográfica realizada revelou que as propostas de estudo de algoritmos numéricos, como os que são tratados neste trabalho, destinam-se ao início de um curso superior ou à fase terminal do Ensino Secundário. Tivemos conhecimento de projetos em que o trabalho elementar em métodos numéricos surge como um tema integrador de aprendizagens, após um período de estudos, como acontece no ensino da matemática no Estado de Victoria, na Austrália [4].

Muito do que lemos sobre estas experiências noutros países poderia ser transportado para a realidade do ensino em Portugal, com as devidas adaptações. Dada a diversidade de cursos destinados a alunos após o 9º ano e ao que se prevê como formação integral de um aluno no final

do Ensino Secundário e tendo ainda em consideração as alterações introduzidas pelo Processo de Bolonha no Ensino Superior, poderá fazer sentido caminhar-se para o aparecimento de um tempo após o Ensino Secundário que permita por um lado uma maior especialização em algumas áreas e por outro lado um tempo bem estruturado que permita aos alunos a mudança de percurso formativo. Na situação de especialização, o estudo de cálculo numérico poderia ter um lugar de destaque em preparação de cursos com uma forte integração da ciência e da prática e também na matemática. Uma mudança destas exigiria alterações relativamente profundas e muito demoradas no sistema educativo em Portugal, condicionadas pela volatilidade de orientações de política educativa. Uma vez que no atual desenho curricular há um leque de disciplinas de opção no 12º ano de escolaridade, uma alteração mais viável poderia ser, por exemplo, a organização de uma destas disciplinas em que os métodos numéricos estejam presentes no âmbito de elaboração de projeto interdisciplinares, numa perspetiva integradora das aprendizagens anteriores em matemática e mesmo noutras disciplinas da área científica. Poderiam também os métodos numéricos integrar outras disciplinas de opção existentes no currículo, como é o caso das aplicações informáticas, em que os algoritmos e a programação têm lugar. No entanto, independentemente da integração destas questões no Ensino Secundário segundo as perspetivas descritas, é possível enriquecer o currículo da disciplina de matemática pela integração de tarefas que se possam explorar no âmbito dos conteúdos curriculares previstos, correspondendo também a metodologias determinantes para a melhor aprendizagem de conteúdos. Neste caso, o trabalho com aspetos elementares que se relacionem com os métodos numéricos, mesmo que não haja a consciência por parte do aluno do método numérico que enquadra a atividade, pode até iniciar-se nos últimos anos do Ensino Básico, como se poderá observar em dois dos exemplos que serão apresentados mais à frente.

Em Portugal, salientam-se em especial as ideias de Sebastião e Silva [6], com quase 40 anos, acerca das vantagens da integração curricular precoce do cálculo de aproximações numéricas num contexto de uma certa crítica ao “ensino tradicional” à época. No momento presente, em que ocorre uma transição entre programas oficiais de matemática, e em que se confrontam visões muito afastadas do que deve ser o ensino da matemática no ensino pré-universitário, as reflexões de Sebastião e Silva poderão ter ainda atualidade. Note-se que vigoram em simultâneo dois programas no Ensino Básico, e a partir do próximo ano letivo (2015/2016) o mesmo acontecerá no Ensino Secundário, até que de, forma gradual, fique apenas um único programa oficial em aplicação.

De [6] transcrevemos algumas frases de Sebastião e Silva para se perceber melhor o que se disse sobre a sua visão do que não deve e do que deve ser o ensino da matemática, em particular, no que diz respeito à integração curricular de conteúdos relativos à utilização de métodos numéricos.

“(...) é essencial que um aluno consiga, ele próprio, sem ajuda, resolver exercícios pela primeira vez. Todo o problema novo, com interesse, tem uma ideia chave, um abre-te Sésamo que ilumina o espírito de súbita alegria (...)” [6, p. 8]

“O nosso ensino tradicional não enferma unicamente de fraca (e quantas vezes nula) insistência em demonstrações, e de insuficiente rigor lógico: peca também por ausência de contacto com o húmus da intuição e com a realidade concreta.” [6, p. 8]

“A crítica dos fundamentos da matemática, iniciada no século passado, conduziu a esse grau de rigor lógico, cuja necessidade se impunha; mas criou ao mesmo tempo um estado de espírito favorável a atitudes rígidas, demasiado platónicas. Seguiu-se uma reacção, por vezes também excessiva, mas em parte salutar, dos chamados ‘matemáticos empiristas.’” [6, p. 9]

“Um ensino que não estimule o espírito e que, pelo contrário, o obstrua com as clássicas matérias para exame, só contribui para produzir máquinas em vez de homens. E não é assim que se curam os males de que está sofrendo o mundo.” [6, p. 11]

No mesmo livro, Sebastião e Silva refere-se então mais especificamente aos aspetos do cálculo numérico, umas vezes de forma mais explícita do que outras:

“Os cálculos exigidos por este processo são laboriosos, mas quando se dispõe de um bom computador, podem ser efetuados rapidamente.(...)”

Entretanto, convém não esquecer este pormenor: do ponto de vista pedagógico, é sempre importante que o aluno conheça, pelo menos, um processo de cálculo, mesmo que não seja o mais expedito.” [6, p. 46]

“Quanto a erros relativos, bastará que o aluno adquira a noção. Será interessante dizer-lhe, a propósito, que os melhores computadores analógicos permitem uma aproximação da ordem de 0,05 %, o que já pode ser considerado muito bom para certos fins. Também haverá interesse em que o aluno aprenda, de modo informal, que o desvio relativo do produto é aproximadamente igual à soma dos desvios relativos dos fatores, etc.” [6, p. 56]

“(...) coloca o aluno imediatamente em contacto com a ideia dos métodos de aproximação, que domina toda a análise numérica moderna, ligada ao uso de computadores. Constitui, por isso também, uma excelente motivação concreta para a introdução do conceito de convergência duma sucessão. O aluno sente que tal conceito é algo de real e de importante, que interessa estudar a fundo. Convém, pois, dedicar um interesse especial ao referido assunto, fazendo-o surgir e desenvolver-se de modo acentuadamente heurístico.” [6, p. 56]

“(...) o aluno pressente que os valores aproximados x_1, x_2, \dots são cada vez mais próximos de α , isto é, tem a intuição de que a sucessão assim definida converge para α . (...)

A demonstração rigorosa só pode ser dada por meio da teoria dos limites, que tem assim, no estudo anterior, uma boa motivação” [6, p. 58]

“Os processos de recorrência (baseados no princípio da indução matemática, que depois será estudado em pormenor) constituem um dos muitos assuntos da matemática que têm sido postos na ordem do dia pelos computadores.” [6, p. 58]

“A teoria dos limites de sucessões, (...) em estreita ligação com o cálculo numérico aproximado, estabelece desde logo uma síntese da teoria com a prática, (...) Escusado será acentuar quanto esta orientação deverá contribuir para despertar o interesse do aluno, que reage quase sempre com desagrado ao aspeto exclusivamente teórico e abstrato de uma teoria dos limites dada a priori, sem qualquer motivação. E note-se que, ao estímulo prático-intuitivo, se segue depois uma estruturação lógica perfeitamente rigorosa.(...)” [6, p. 67]

“Estaremos assim, mais uma vez., a proceder à semelhança do que se faz em investigação. Se porventura Newton, Leibniz, Euler, Lagrange, Laplace e outros mais tivessem ficado à espera de uma demonstração rigorosa dos seus resultados antes de os publicarem, não teriam sido possíveis os enormes progressos que, desde então até hoje, se têm realizado em matemática e nas ciências afins.” [6, p. 70]

A frase seguinte, embora à data mais direcionada para o Ensino Superior, poderá ser considerada adequada atualmente para o Ensino Secundário.

“Os (...) exercícios conduzem, de modo natural, a redescobrir um método geral para o cálculo numérico de todas as raízes de uma equação algébrica (...). Os fundamentos deste método, relativamente elementar, são acessíveis a qualquer bom aluno que, porventura, sinta curiosidade pelo assunto. Por outro lado, as dificuldades de cálculo numérico inerentes ao método estão hoje

em grande parte removidas pelos computadores eletrónicos. Por isso não resistimos à tentação de o inserir no texto, a título facultativo, com exemplos numéricos, pensando sobretudo numa das várias incongruências que se verificam no ensino universitário da matemática: os alunos aprendem aí teorias, mais ou menos profundas, relativas a equações algébricas; mas se alguém lhes perguntar como se calculam todas as raízes de uma dada equação algébrica, de grau arbitrário, com a aproximação que se queira, terão de reconhecer que não sabem. Isto dá bem a nota de quanto o ensino tradicional da matemática tem sido afastado da realidade.” [6, p. 70]

A propósito de ter mostrado como um problema prático envolvendo o cálculo de um integral pode ajudar os alunos a compreender as “linhas mestras” da integração, Sebastião e Silva advoga a “verdadeira prática”, segundo ele muito diferente da “pseudoprática” dos cursos tradicionais (secundários ou universitários) que é feita “à base de inúmeras receitas que na maioria dos casos nunca virão a ser aplicados”:

“Haveria muitíssimo a lucrar em que o ensino destes assuntos fosse normalmente orientado a partir de centros de interesse como o anterior - e tanto quanto possível laboratorial, isto é, baseada no uso de computadores existentes nas próprias escolas ou fora destas, em laboratórios de cálculo.” [6, p. 89]

Ao longo do tempo podem encontrar-se em Portugal exemplos pontuais de tratamento de métodos numéricos antes do ensino superior. A título meramente ilustrativo, e por serem do nosso conhecimento, refira-se o tratamento do método de Newton-Raphson e das bisseções na disciplina de Cálculo Numérico do 12º ano que integrava um curso de oferta própria do Colégio S. Gonçalo, em Amarante, e que era ministrado por professores de informática, do método do ponto fixo (de forma elementar) no âmbito de aulas de substituição para que era chamado um professor da Escola Secundária Virgílio Ferreira, em Lisboa, e do método da bisseção, ao nível do 12º ano da disciplina de Matemática A, a propósito do tratamento do teorema de Bolzano. Acerca da utilização do método numérico do ponto fixo na Escola Secundária Virgílio Ferreira, o Professor José Paulo Viana afirmou, em entrevista, que tratava este assunto nas aulas em que substituíra professores em falta, em turmas do 3º ciclo do ensino básico. Utilizava uma “brincadeira” com a calculadora que mais não era do que pedir aos alunos para introduzirem um número qualquer, afinal a estimativa inicial para uma raiz da equação $x = \varphi(x)$, e ao premir repetidamente a tecla apropriada fazia funcionar a função iteradora φ . Em muitos casos o número para o qual tendia a

sucessão de números produzida (e que finalmente se repetia na máquina sem alteração) era o mesmo, independentemente do ponto de partida. Esta forma simples de produzir na calculadora sucessões de números diferentes mas convergindo para o mesmo ponto intrigava os alunos que manifestavam um grande interesse por saber qual era a explicação do “truque mágico”.

4.3. TAREFAS PARA A SALA DE AULA

Um dos aspetos mais elementares do cálculo numérico, ao mesmo tempo estruturante, tem que ver com as questões das aproximações e dos vários tipos de erros, em particular os erros (de arredondamento) que resultam da forma de representar os números na máquina. É importante que os alunos estejam preparados para compreender as limitações tecnológicas que estão na origem destes erros de arredondamento e as suas consequências. Por exemplo, deve ser chamada a atenção dos alunos para o facto de que, na calculadora, o resultado da soma $1 + x$ pode ser igual a 1 sem que o valor de x seja 0.

Em termos curriculares, questões muito elementares relativas a aproximação e arredondamento surgem logo entre o 1º e o 2º Ciclos do Ensino Básico, como se pode ver em [32, p. 12 e 15], por exemplo. Ao nível do 8º ano, por exemplo, é previsto o trabalho com dízimas e representação na reta numérica [32, p. 22] e no 9º ano é explícita a necessidade de trabalhar com valores aproximados de resultados de operações, com destaque para problemas envolvendo aproximações de medidas de grandezas [32, p. 24]. Há assim lugar ao tratamento destas questões de forma curricular mesmo antes da entrada no Ensino Secundário e a oportunidade de referir alguns aspetos não explicitamente previstos como a questão do erro absoluto e do erro relativo, da diferença e do interesse de cada um destes, apesar de no caderno de apoio a [32] apenas haver algumas referências a majoração de erros no sentido de erro absoluto. O erro relativo, mesmo que não surja em [32] ou nos respetivos cadernos de apoio, pode ser abordado com muito interesse por causa das questões da medida e da utilização de instrumentos de desenho nas construções. Se, por exemplo, um professor pretende que o aluno conjecture sobre as relações entre medidas em polígonos, poderá sugerir que não construa polígonos demasiadamente pequenos. A explicação de que o mesmo erro (absoluto) na utilização do instrumento de desenho pode provocar erros mais significativos numa figura pequena do que numa maior parece-nos uma forma bastante adequada de fazer aparecer o erro relativo e o seu interesse.

Ao nível do 3º ciclo do Ensino Básico são trabalhadas as sucessões [32, p. 19]. Embora a definição por recorrência só surja no Ensino Secundário [33, p. 18], a familiarização com a ideia de

recorrência no 3º ciclo poderá melhorar a compreensão dos assuntos a estudar mais tarde, incluindo os próprios limites de sucessões.

Os métodos iterativos já não se devem trabalhar ao nível do Ensino Básico, nem mesmo no início do Ensino Secundário, pois são assuntos de dificuldade superior para se tratarem com objetivo de apreciar as suas potencialidades, embora o método do ponto fixo, por exemplo, possa assumir desde o 3º ciclo um lugar de destaque, sendo a abordagem feita por José Paulo Viana um exemplo viável. É possível elaborar tarefas que possam ser utilizáveis ao nível do 7º ano no tema “equações”, porventura com mais interesse no 8º ano no tratamento de conteúdos diversos, como as “equações”, novamente, ou os “gráficos de funções afins” ou a “resolução de sistemas de duas equações e duas incógnitas”. Obviamente que neste trabalho elementar a expressão “método do ponto fixo” pode não estar presente. A este propósito, o início das explorações referidas em [34] dá uma ideia e um enquadramento interessante sobre o que se poderá fazer no Ensino Secundário e até onde é possível um tratamento mais elementar. É aí referido que os estudantes desenvolvem numa folha de cálculo o algoritmo do método do ponto fixo com função iteradora da forma $\varphi(x) = ax + b$, começando por dois parâmetros reais a e b desta família de funções e percorrem um caminho de descoberta de condições de convergência e da relação entre parâmetros para concluírem sobre o valor para o qual converge. No mesmo documento é mencionada a possibilidade de um trabalho análogo, mas um pouco mais complexo, com funções iteradoras da forma $\varphi(x) = ax(1 - x)$. Estas questões poderão ser colocadas aos alunos com níveis de aprofundamento bastante diversos, de acordo com a estratégia educativa do professor, podendo ser retomados e aprofundados *à posteriori*.

Um método como o de Newton-Raphson exige o conhecimento da função derivada para ser desenvolvido numericamente, pelo que é aconselhável tratá-lo a partir do 11º ano [33, p. 19]. A interpretação geométrica deste método está intimamente ligada ao estudo de equações de retas tangentes ao gráfico de uma função e o teorema (do valor médio) de Lagrange regressa aos programas de matemática do Ensino Secundário que entrarão em vigor em 2014/2015 [33, pp. 16, 19 e 40], havendo com isso mais uma oportunidade de utilização do método de Newton-Raphson no Ensino Secundário.

O método da bisseção tem mais interesse a par do tratamento do teorema de Bolzano-Cauchy no 12º ano [33, p. 24], depois de trabalhado o conceito de função contínua no 11º ano [33, p. 19], embora nada seja referido explicitamente em [33].

Muitas das tarefas que se possam conceber poderão passar pela utilização de uma máquina de calcular, o que é compatível com os programas oficiais de matemática, claramente nos do secundário [33] e mesmo nos do básico [32]. No programa oficial de matemática do Ensino Básico [32, p. 28 e 29] é referida a possibilidade das calculadoras serem utilizadas em situações pontuais de resolução de problemas que envolvam, por exemplo, um elevado número de cálculos ou a utilização de valores aproximados, apesar do alerta para que não seja comprometida, com a sua utilização, a aquisição de procedimentos e o treino do cálculo mental. No Ensino Secundário [33] prevê-se também a possibilidade de utilização da calculadora gráfica, aqui de forma mais alargada que no Ensino Básico, chamando a atenção para a importância de alertar os alunos para as limitações da máquina de calcular (p. 28) ou a obtenção de valores aproximados de soluções de equações (p. 29).

Nas páginas seguintes apresentam-se algumas sugestões de tarefas com uma estrutura próxima da que poderá ser usada em sala de aula, concluindo-se com algumas observações gerais em relação ao conjunto das tarefas. São privilegiados conteúdos do Ensino Secundário, sobretudo dos últimos anos (11º e 12º), embora se possam perspetivar casos pontuais de parcelas das tarefas que podem ser abordadas nos últimos anos do 3º ciclo do Ensino Básico.

EXEMPLOS DE TAREFAS PARA A SALA DE AULA

TAREFA 1 – Vamos conhecer melhor a tua máquina calculadora

Verifica se $x = 0.0000000000000001$ (as primeiras 14 casas decimais são iguais a zero) leva a que $1 + x$ produza o resultado 1.

Repete o cálculo da soma variando o número de casas decimais iguais a zero no valor de x e determina o menor número positivo que seja uma potência de 10, com expoente negativo, tal que $1 + x$ dá um resultado maior do que 1 na tua calculadora (é natural que outro modelo de calculadora possa levar a um resultado diferente).

TAREFA 2 – Calcular o inverso com as teclas da divisão e do expoente avariadas

Nos primeiros modelos de computadores digitais a divisão não era efetuada por hardware mas sim por software. Assim, a divisão de a por b , com a e b positivos, era efetuada multiplicando a pelo inverso de b , pelo que o problema se transferia para o inverso de um número.

1. Uma vez que $\frac{1}{b}$ é a solução da equação $b - \frac{1}{x} = 0$, mostra que o método de Newton-

Raphson fornece a seguinte fórmula iterativa para o cálculo de $\frac{1}{b}$:

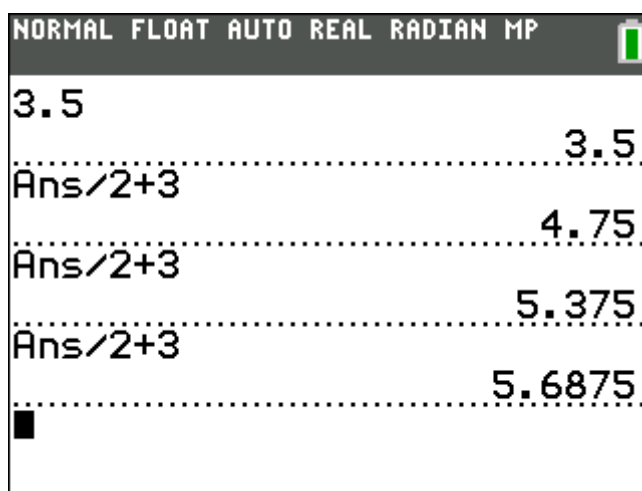
$$x_{n+1} = x_n(2 - bx_n), \quad k = 0, 1, 2, \dots$$

2. Calcula, por este processo o valor de $\frac{1}{7}$ tão exatamente quanto a tua máquina de calcular permitir começando com a aproximação inicial 0.1 e observa que nas últimas iterações, o número de algarismos corretos praticamente duplica em cada iteração.
3. Tenta calcular o valor de $\frac{1}{7}$ partindo de 0.3. O que acontece? Experimenta outros valores iniciais e tenta encontrar um padrão no comportamento das iterações.

TAREFA 3

- 1 – Prime as teclas necessárias até aparecer um número à tua escolha no visor da máquina e prime a tecla **enter** (noutros modelos pode ser **=**);
- 2 – Depois, prime sucessivamente as teclas **÷****2****+****3** e observa o resultado;
- 3 – De seguida, prime a tecla **enter** (ou **=**) sucessivamente e observa o que sucede;
- 4 – Finalmente, regista o número inicial e o que observas depois de premir **enter** (ou **=**) algumas vezes mais.
- 5 – Repete o procedimento, iniciando com outros números.

Depois de alguns registos repete o processo e, no passo 2, prime sucessivamente **×****2****-****9**.
Repete com outros números à tua escolha.

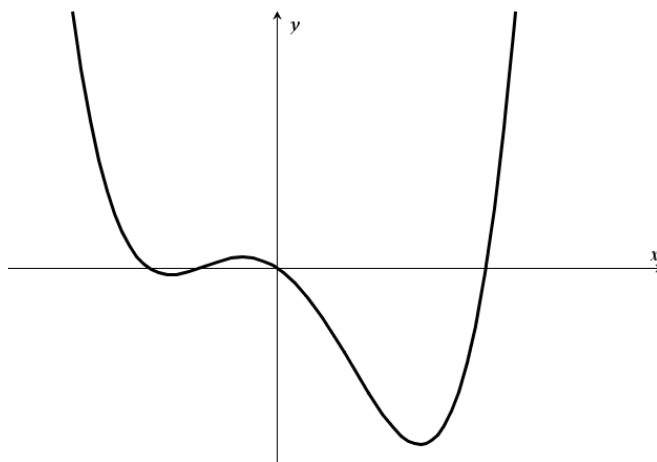


Que poderás concluir?

TAREFA 4

Considera a função f definida por $f(x) = x^4 - 2x^2 - x$, representada graficamente mais abaixo.

Sabe-se que esta função tem um único zero positivo e supõe-se que apenas há disponível uma máquina de calcular com as quatro operações elementares, papel e lápis.



1. Mostra que esse zero positivo se localiza no intervalo $[1,2]$.
2. Se se considerar o valor central do intervalo $[1,2]$, quanto pode diferir, no máximo, este valor do zero da função? E se tivesses considerado outro valor no intervalo?
3. Em qual dos intervalos obtidos pela bisseção do anterior se localiza o zero, em $[1,1.5]$ ou em $[1.5,2]$?
4. Qual é agora o erro máximo que se pode cometer ao considerar o valor central do intervalo selecionado na alínea anterior como zero de f ? E se o valor considerado tivesse sido outro?
5. Continuando o processo de bisseções sucessivas do intervalo onde o zero se localiza, determina um intervalo de amplitude inferior a 0.001 e que contem o zero procurado. Conta o número de bisseções que foram necessárias para obter esta aproximação.
6. Sem obter os intervalos sucessivos, determina o número de bisseções que são necessárias para obter um intervalo de amplitude inferior a 0.00001.
7. Será o zero positivo de f o número de ouro?

TAREFA 5

Considera a função g definida por $g(x) = x^2 - 2$ e a sucessão (u_n) definida por:

$$\begin{cases} u_1 = 3 \\ u_{n+1} = u_n - \frac{g(u_n)}{g'(u_n)} \end{cases}$$

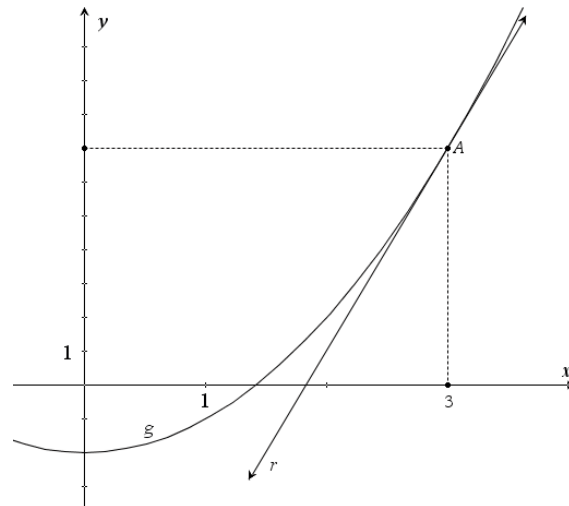
1. Calcula o 2º termo da sucessão.

2. Mostra que (u_n) se pode definir por $\begin{cases} u_1 = 3 \\ u_{n+1} = \frac{u_n^2 + 2}{2u_n} \end{cases}$

3. Completa a tabela com alguns termos de (u_n) . Podes obter um valor aproximado do 4º termo com 3 casas decimais, por arredondamento.

n	1	2	3	4
u_n	3			

4. Considera agora o gráfico da função g , cuja representação gráfica podes observar na figura ao lado, na qual também se observa, no mesmo referencial, a reta r , tangente ao gráfico de g no ponto A, de abcissa 3.



4.1. Determina uma equação da reta r e mostra que interseja o eixo das abcissas no ponto $\frac{11}{6}$ (P_1).

4.2. Considera, pelo ponto B, do gráfico de f com abcissa $\frac{11}{6}$, uma nova reta s , tangente ao gráfico de g . Qual será agora a abcissa do ponto P_2 , de interseção da reta s com o eixo das abcissas?

5. Podes observar que as abcissas dos pontos P_1 e P_2 coincidem com os dois primeiros termos da sucessão (u_n) , respetivamente.

Mostra que este facto não é simples coincidência, ou seja, mostra que os termos da sucessão (u_n) , obtidos por recorrência, são as abcissas das sucessivas retas tangentes consideradas como em 4, após ter a reta r .

6. Estuda a evolução dos termos da sucessão (u_n) e relaciona com a função g .

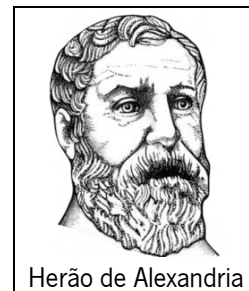
7. Será que tens um método geral para determinar zeros de funções ou para resolver equações?

Experimenta com outras equações.

TAREFA 6

Atualmente é muito fácil obter valores com excelente aproximação para a raiz quadrada de um número com um simples toque numa máquina de calcular. Mas antes da Era dos computadores e das máquinas de calcular do tipo das que hoje utilizamos, chegar à raiz quadrada era uma tarefa mais difícil.

Na sua prática de engenharia, Herão de Alexandria (geómetra e engenheiro grego, que viveu muito provavelmente no século I da Era cristã), tinha de resolver muitos problemas envolvendo raízes quadradas e para tal utilizava um método, que não terá inventado mas que hoje em dia se conhece pelo seu nome (também designado por método babilónico).



Basicamente, para calcular a raiz quadrada, digamos r , de um número real positivo x , procede-se da maneira seguinte:

- 1º - Procura-se um valor aproximado, digamos apróx para a raiz (isto não é difícil, como verás no exemplo);
- 2º - Calcula-se a média aritmética dos valores apróx e $\frac{N}{\text{apróx}}$, obtendo-se assim um novo valor apróx ;
- 3º - Repete-se o procedimento anterior tantas vezes quantas as necessárias até obter a raiz quadrada com a precisão desejada.

Observando o exemplo seguinte, este processo poderá ser compreendido.

Vamos procurar a raiz quadrada de 12. Podemos escolher 3.5 como primeira aproximação, pois como $3^2 = 9$ e $4^2 = 16$, a raiz quadrada está entre 3 e 4.

A segunda aproximação é $\frac{1}{2} \left(3.5 + \frac{12}{3.5} \right) = 3.4642857 \dots$

Esta 2ª aproximação, fácil de obter na calculadora, apresenta já 4 dígitos significativos corretos, o que poderá ser em muitas situações uma excelente aproximação.

1. Encontra aproximações das raízes quadradas de 14, de 26, de 35 e de 87, utilizando o método de Herão. Utiliza duas aproximações e arredonda às centésimas. Compara os resultados com os valores das raízes quadradas obtidas nas máquinas de calcular.
2. Para perceberes porque é que o método funciona, pensa no seguinte: se o valor inicial apróx for superior ao valor da raiz r , o que se pode dizer de $\frac{N}{\text{apróx}}$?
E se, pelo contrário, o valor inicial de apróx for inferior ao valor da raiz?
3. Se já conheces o método de Newton-Raphson, usa-o para calcular uma aproximação da raiz da equação $x^2 = 12$ a partir da aproximação inicial 3.5. Compara o valor obtido com o que está no exemplo dado antes. O que conclusis?

TAREFA 7

Nos meios bancários e financeiros é conhecida a equação seguinte:

$$E = \frac{P}{t} \left[1 - \frac{1}{(1+t)^n} \right],$$

onde E representa o valor de um empréstimo concedido pela instituição financeira, P o valor (fixo) de cada amortização periódica, n o número de pagamentos (amortização da dívida) e t a taxa de juro (relativa ao período).

Supõe que alguém vai precisar de um empréstimo no valor de 100 000 euros e pretende pagar a dívida em 30 anos com uma prestação mensal fixa de 450 euros. Para saber qual é o valor da taxa de juro mensal que deve negociar com o banco, resolve a equação

$$100000 = \frac{450}{t} \left[1 - \frac{1}{(1+t)^{30 \times 12}} \right],$$

usando um método à tua escolha. Expressa o resultado em termos percentuais

TAREFA 8

Uma das raízes da equação $x^2 + bx + c = 0$ é dada por

$$r = \frac{-b + \sqrt{b^2 - 4c}}{2} \quad (1)$$

ou

$$r = \frac{2c}{-b - \sqrt{b^2 - 4c}} \quad (2)$$

Como resulta de multiplicar o numerador e denominador de (1) por $-b - \sqrt{b^2 - 4c}$.

Na máquina de calcular usa as fórmulas (1) e (2) para calcular o valor de r no caso de ser $b = 2^{21}$ e $c = 2^{-14}$.

Qual dos valores obtidos aproxima melhor o valor exato da raiz r ?

Tenta compreender a causa principal do erro na pior aproximação.

BREVE COMENTÁRIO ÀS TAREFAS APRESENTADAS

A primeira tarefa apresentada anteriormente e a oitava (em parte), a propósito da possibilidade de integração curricular, e a experiência referida no início do segundo capítulo são exemplos de como alertar os alunos para algumas limitações das máquinas. As tarefas seguintes permitem diversidade de aprofundamento, tornando-as adaptáveis de acordo com o interesse pedagógico do professor. Aliás, dado o carácter aberto com que algumas são apresentadas, permitindo um percurso de trabalho investigativo, poderão estabelecer-se etapas de complexidade cognitiva crescente que podem ser percorridos em parte ou na totalidade, num momento ou em vários momentos, por todos os alunos ou apenas parte deles. A título de exemplo refira-se que a tarefa **3** poderá iniciar-se para todos os alunos, e desde cedo, com os procedimentos referidos de 1 a 3, suscitando a observação da surpreendente regularidade, podendo também cada aluno passar à etapa do registo. A descoberta de outras regras e a obtenção de diferentes resultados seria o passo seguinte em termos de complexidade, evoluindo para a exploração analítica, obtenção de regras com objetivo de ponto fixo, exploração geométrica e interpretação das explorações sob esse ponto de vista, conjecturar sobre condições de convergência, justificar com algum nível de rigor essas condições de convergência e a implementação mais alargada e consciente do método do ponto fixo para resolver equações. Esta tarefa é por isso oportuna em diversas ocasiões, desde a aplicação elementar no âmbito da resolução de equações e sistemas de equações no 8º ano de escolaridade. As tarefas **2**, **5** e **6** utilizam uma base comum, o método de Newton-Raphson, para cobrirem diversas possibilidades de utilização e diferentes modos de trabalho com os alunos. A tarefa **6** pode levar uma motivação especial aos alunos pela inclusão da história da matemática na determinação aproximada da raiz quadrada de um número. Até à conclusão do 1º item pode haver quem a considere uma tarefa bastante apropriada para alunos mais novos. Na globalidade, a tarefa permite uma contextualização histórica e a direta relação entre dois métodos numéricos de resolução de equações separados no tempo por muitos séculos, e ainda a compreensão dos procedimentos referidos na primeira secção do segundo capítulo a propósito da utilização da calculadora mecânica Curta para obter a raiz quadrada de um número. A tarefa **5** permite um percurso de obtenção da fórmula de Newton-Raphson associada à interpretação geométrica, com ênfase nas sucessões e nas derivadas, relativas ao 11º ano de escolaridade. A tarefa **2**, embora de formulação mais simples, necessita já do conhecimento da fórmula de Newton-Raphson e permite um aprofundamento maior por causa do efeito “inesperado” por se considerarem aproximações iniciais diferentes e permitirá análises gráficas e eventualmente aproximação às

condições de convergência, num contexto de cálculo de um simples inverso de um número com ligeiro enquadramento histórico que favorece a motivação. É uma tarefa de carácter claramente investigativo, que ao considerar uma sequência temporal teria de vir após a tarefa **5**, ou análoga, e que mais uma vez mobiliza alguns assuntos do Ensino Secundário. O método da bissecção faz também parte do leque de tarefas, tarefa **4**, permitindo revelar ao aluno, que não necessita das ferramentas necessárias ao método de Newton-Raphson, uma forma de resolução de equações que, com a mesma dificuldade, ou facilidade, pode levar à resolução de equações que a via analítica não permite. Dado o aspeto intuitivo do método, ele pode surgir desde que há uma representação gráfica de funções polinomiais, logo no Ensino Básico, mas por necessitar de conceitos como a continuidade ou o teorema de Bolzano-Cauchy, abordados no final do Ensino Secundário, poderá ser tratada com maior vantagem por essa altura, a menos que alguma oportunidade surja anteriormente. A última questão desta tarefa **4** é um pouco distinta, de carácter analítico, que constitui uma situação problemática que pode levar o aluno a pesquisar sobre o número de ouro, sobre uma equação do 2º grau que o tem como solução e a sua relação com o retângulo de ouro, onde a regra de Ruffini tem oportunidade de ser revisitada. É uma tarefa integradora, como muitas que podem ser realizadas com o cálculo numérico, o que pode favorecer a compreensão da matemática e ajudar à preparação para ambiente de exame num contexto novo que é o facto do exame do 12º ano de matemática. A partir de conteúdos desde o 10º ano de escolaridade. A tarefa **7** tem uma particularidade que a distingue das restantes, que é a possibilidade de uma comparação entre diferentes métodos num contexto real e atual, e por isso previsivelmente motivador, em que uma equação de resolução impossível por via analítica suscita a necessidade de uma resolução por aproximação em que não é referido à partida o método a utilizar. A tarefa **8** traz também a problemática da representação de números em máquinas e enfatiza a problemática da propagação dos erros nas operações, com destaque para o cancelamento subtrativo. A última sugestão de trabalho poderá não ser aconselhável para alunos do Ensino Secundário, mas não deixa de se colocar porque haverá um ou outro aluno que poderá interessar-se por um aprofundamento a este nível. O trabalho que aí possa ser feito deverá ser suscitado pelo aluno e não por indicação do professor.

Ficam assim um conjunto de ideias para levar à prática e com isso perspetivar em contexto a utilidade da inclusão das questões do cálculo numérico na melhoria da aprendizagem dos alunos do Ensino Secundário, que estamos convencidos acontecerá.

BIBLIOGRAFIA

- [1] R. A. Gonçalves, 2001. [Online]. Available: <http://espmat.no.sapo.pt/intercambio/conferencia.html>. [Acedido em 19 10 2014].
- [2] P. W. F. & R. D. West, "Math Archives," 1996-2001. [Online]. Available: <http://archives.math.utk.edu/ICTCM/VOL15/W005/paper.pdf>. [Acedido em 01 02 2014].
- [3] "Victorian Curriculum and Assessment Authority, Victorian Certificate of Educational Study design," Victoria, Austrália, 2010.
- [4] D. L.-L. & M. Evans, Numerical Methods in Senior Secondary Mathematics, Austrália: Shaping Australia, 2000.
- [5] A. J. G. d. Gouveia, "Colégio S. Gonçalo," 2010/2012. [Online]. Available: <http://www.csgnet.org/documentosONLINE/apontamentos/ecn/ecn5.pdf>. [Acedido em 01 02 2014].
- [6] J. S. e. Silva, Guia para utilização do compêndio de Matemática (2º e 3º volumes), Lisboa: GEP, ME, 1977.
- [7] P. Teixeira, Funções: Matemática - 10º ano, Departamento do Ensino Secundário, Lisboa: Editorial do Ministério da Educação, 1997.
- [8] D. d. S. d. I. d. U. d. Minho, "Museu Virtual da Informática," 08 2001. [Online]. Available: <http://piano.dsi.uminho.pt/museuv/1820thollerith.html>. [Acedido em 12 10 2014].
- [9] A. Logioia, "Curtamania," 13 8 2014. [Online]. Available: http://www.curtamania.com/curta/code/curta_manuais.html. [Acedido em 14 10 2014].
- [10] R. Bensene, "The old calculator web museum," 27 4 2014. [Online]. Available: <http://www.oldcalculatormuseum.com/pocketronic.html>. [Acedido em 25 9 2014].
- [11] IEEE, "IEEE - Advancing Technology for Humanity," IEEE, 2014. [Online]. Available: http://www.ieee.org/about/news/2009/14april_1.html. [Acedido em 8 10 2014].
- [12] G. Melo, "Gilberto Melo - Engenharia Jurídica," [Online]. Available: <http://www.gilbertomelo.com.br/rss/26-calculos/1457-primeira-calculadora-eletronica-completa-40-anos>. [Acedido em 25 9 2014].

- [13] M. R. Valença, *Análise Numérica*, Lisboa: Universidade Aberta, 1996.
- [14] IEEE, "IEEE - Advancing Technology for Humanity," [Online]. Available: <http://standards.ieee.org/findstds/standard/754-2008.html>. [Acedido em 01 02 2014].
- [15] P. Gomes, Universidade do Minho, 2005, Tese de Mestrado, <http://hdl.handle.net/1822/6303>.
- [16] T. Sauer, *Numerica Analysis - Second Edition*, Boston, USA: Pearson Education, Inc., 2012.
- [17] M. R. Valença, *Métodos Numéricos*, Braga: Livraria Minho, 1993.
- [18] D. N. Arnold, "Institute of Mathematics and its Applications," University of Minnesota - USA, 8 9 2014. [Online]. Available: <https://www.ima.umn.edu/~arnold/disasters/patriot.html>. [Acedido em 8 10 2014].
- [19] H. Pina, *Métodos Numéricos*, Lisboa: McGraw-Hill, 1995.
- [20] K. E. Atkinson, *An introduction to Numerical Analysis - 2nd ed.*, United States: John Wiley & Sons, 1988.
- [21] M. R. Valença, *Métodos Numéricos*, Instituto nacional de Investigação Científica, 1998.
- [22] R. L. Burden & J. D. Faires, *Numerical Analysis*, 9th Edition, Brooks/Cole Publishing Company: Cengage Learning, 2010.
- [23] "isaw.nyu.edu," [Online].
- [24] "www.maa.org," [Online].
- [25] O. J. Brison, *Teoria de Galois*, 4ª edição, Lisboa: Dep. Matemática da Faculdade de Ciências de Lisboa, 2003.
- [26] F. Saleri & A. Quarteroni, *Cálculo Científico com Matlab e Octave*, Milão - Itália: Springer, 2007.
- [27] M. A. Nordgaard, *A Historical Survey of Algebraic Methods of Approximating the Roots of Numerical Higher Equations up to the year 1819*, New York City: Teachers College, Columbia University, 1922.
- [28] P. Dawkins, "Paul's Online Math Notes," [Online]. Available: <http://tutorial.math.lamar.edu/Classes/Calcl/NewtonsMethod.aspx>. [Acedido em 13 08 2014].
- [29] "Numberempire," 7 outubro 2011. [Online]. Available: <http://pt.numberempire.com/equationsolver.php/equationsolver.php>. [Acedido em 30 agosto 2014].

-
- [30] M. P. J. Carpentier, "Análise Numérica," Instituto Superior Técnico - UTL, Lisboa, 1993.
- [31] J. C. e Silva, "Understanding Exponential Growth with Technology," ICME 7, Ohio - US, 1994.
- [32] Oliveira, Filipe *et al*, "Programas e Metas Curriculares - Matemática - Ensino Básico," DGE, Lisboa, 2013.
- [33] Oliveira, Filipe *et al*, "Programas e Metas Curriculares de Matemática A," DGE, Lisboa, 2014.
- [34] J. A. Walsh, "Iteration in first semester Calculus," ICTM 7, Ohio - US, 1994.

