

As Múltiplas faces da Derivada I: Newton vs. Leibniz

The Multiple faces of the Derivative I: Newton vs. Leibniz

Wilson Castro Ferreira Junior*¹

¹Departamento de Matemática Aplicada, UNICAMP-IMECC- Campinas – São Paulo – Brasil

Resumo

O conceito de derivada é sutil e representa uma das barreiras mais sérias dentre as várias que dificultam o ensino e o aprendizado, do Cálculo. Um indício forte da causa desta dificuldade pode ser diagnosticado quando observamos que a organização do ensino desta disciplina subverte totalmente a sua evolução histórica em mais de mil anos, invertendo a ordem entre a introdução do conceito de derivada e do conceito de integral, que são os dois limites basilares da disciplina. É de se imaginar a razão pela qual mentes tão poderosas como as de Arquimedes e Galileu, para citar apenas duas dentre as mais brilhantes que já existiram, não atinaram com o conceito de derivada durante este enorme interlúdio. O fato é que o conceito de derivada é notoriamente difícil, mesmo para mentes privilegiadas, por se tratar de um limite singular ($\frac{0}{0}$)! O objetivo deste artigo é apresentar o conceito de derivada, apoiando-se na definição intuitiva de G.W. Leibniz que enfatiza o conceito de "aproximação" em lugar do conceito de limite singular de Newton deixando de lado completamente a (importante) definição "com ϵ 's e δ 's" introduzida por A.L. Cauchy no início do século XIX que é religiosamente repetida "ad nauseam" por virtualmente todos os textos do assunto. Em apoio ao argumento apresentam-se diversos exemplos "concretos", nem sempre encontrados na literatura pedagógica, que ilustram as diversas "faces" da derivada o que, se espera, venha em auxílio daqueles que dão início ao estudo do Cálculo com o intuito de entendê-lo.

Palavras-chave: Cálculo, Derivada de Leibniz, História do Cálculo.

Abstract

In a typical first course of Calculus the students are usually introduced to the idea of derivative as a point operation many pages, definitions, theorems and exercises prior to the treatment of the idea of integral. This usage is at least 1000 years in the opposite direction to the historical development of the mathematical ideas. One may wonder why such giants as Arquimedes and Galileu to name just two of the most powerful mathematical minds ever to exist, have not caught up with the idea of derivative during that long dormancy. We do not have to reach for a historical argument, ponderous as it might be, to assert that the derivative as a function operation is a notoriously difficult concept to face head on, more so if subjected to the too formal and far from intuitive Cauchy's " $\epsilon - \delta$ " approach to the singular ($\frac{0}{0}$) limit. It is safe to say that a giant like G.W. Leibniz had no trouble to understand Newton's definition but, by all means, he decided (before or after seeing the Newtonian procedure) for an alternative path. The Leibnizian approach to derivative, although mathematically equivalent to Newton's has a quite different motivation as a starting point. Leibniz based his work on a concept of "approximating" general functions by the most simple ones, not from a mere numerical point of view though. What makes Leibniz approach interesting nowadays is that we can repeat it easily without that mysterious aura of his "infinitesimals" and, at the same time, get some huge fringe benefits, like redefining also the idea of limit, making it more intuitive, useful and ready for generalization. The substitute for the Leibnizian infinitesimals are the concept of "scales" and "order" as a way of setting standards for measuring the "rapidity" of approach. A few but illustrious examples of Leibnizian derivatives not usually found in textbooks will help understand this rich concept through the many faces it presents in Mathematics and its Applications.

Keywords: Calculus, Leibnizian Derivative, History of Calculus.

1 Introdução

As disciplinas de Cálculo, em todo o mundo, constituem uma notória e séria barreira acadêmica e intelectual das grades curriculares nos primeiros anos de ensino superior. Esse "fenômeno" não deveria surpreender a quem examinar com cuidado a maneira inconsistente e mesmo estranha como uma das construções mais importantes e abstratas do intelecto humano é apresentada em classe. Dentre os graves problemas metodológicos no ensino desta disciplina, especial relevo deve ser atribuído à bisecular insistência na introdução do fundamental conceito de limite à *la Cauchy*. Repete-se hoje quase que literalmente o mesmo, ainda que genial, argumento inventado por A.L. Cauchy (1789-1857) em 1821 para organizar o descalabro lógico que viciava os resultados do Cálculo de então. Hoje, esse procedimento simplesmente adentra o aluno na obtenção passiva dos "deltas dados os epsilons" para alguns poucos exemplos elementares, o que em quase nada contribui para o seu aprendizado de um conceito tão sutil. Argumenta-se, ilusoriamente, que esse é "*o tratamento rigoroso*" do assunto, restando saber em que sentido e a quem deva ser dirigido. A maioria dos exemplos tratados a partir da definição de Cauchy em textos e classes são limites regulares e triviais, exceção feita a alguns limites singulares ($\frac{0}{0}$) chamados *notáveis* (exemplo de fato notável da confusão do que seja rigor matemático, encontra-se no próprio estudo do limite $\lim_{x \rightarrow 0} \frac{\text{sen } x}{x} = 1$, quando, via de regra, a função seno é apresentada geometricamente como se fosse um objeto matemático ao mesmo tempo em que ela é obviamente obtida como modelo matemático de um objeto Físico, o Círculo. Perde-se assim uma oportunidade excepcional de discutir o importante conceito de modelo matemático, que raramente é mencionado, em troca de um exemplo enganoso de rigor matemático).

O presente artigo se dedica a um outro importante gargalo no aprendizado do Cálculo, decorrente imediato da dificuldade herdada do conceito de limite, a operação derivada.

Em um típico curso desta disciplina, o conceito de *derivada* é, logo de início, apresentado na sua forma mais geral possível como resultado de um limite pontual e singular, $\frac{df}{dx}(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$, ou, pior: $y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$. Para o aluno que mal entendeu a ideia de limite regular, o novo conceito é recebido obscuramente e carregado com todos os tons da mistificação. A bem da verdade, a sua definição "*rigorosa*" é esquecida em pouco tempo, tanto pelos instrutores quanto pelas defesas cognitivas dos alunos, quando o assunto é sorrateiramente substituído pela apresentação receituária da *operação derivada*. Na maioria dos textos, a passagem do conceito de *derivada* como um limite pontual para o conceito de operação funcional de *derivada* se dá

sem maiores cerimônias, perdendo-se aí uma outra preciosa oportunidade de enfatizar um dos aspectos mais fundamentais e cruciais na compreensão de todo o Cálculo, tanto como teoria como método matemático. Ao fim, as sutilezas dos conceitos de derivação/derivada são varridos para debaixo do tapete, atitude em que seria oportuno lembrar o histórico e bem humorado encorajamento de d'Alembert aos seus alunos: "*Moçada, vamos em frente, que o hábito haverá de convencê-los, eventualmente*".

A prevalência da atual metodologia para o ensino do Cálculo é um fenômeno em tudo semelhante à dominação de certas tecnologias modernas que, apesar de serem muito inferiores às algumas concorrentes, acabaram por serem adotadas em decorrência de circunstâncias históricas e do poder econômico que as tornaram tão sofisticadas e caras, mas razoavelmente eficientes que, hoje, seu abandono seria inviável (Exemplo desse processo foi a dominação do sistema de gravação VHS sobre o superior sistema Betamax e, por outros motivos, a desconcertante disposição das letras de um teclado que nada tem de razoável. O desengonçado motor a pistão que transforma movimentos bruscos de vai-e-vem em um movimento rotacional, é talvez o exemplo mais impressionante).

No que diz respeito ao conceito de derivada, uma das inconsistências mais graves na metodologia adotada para o ensino de Cálculo é a sua sistemática apresentação antes do conceito de integral, o que deveria escandalizar a muitos quando sequer é notada. Esse procedimento contraria totalmente a ordem do desenvolvimento histórico da Matemática, o que por si só já deveria ser um estridente sinal de alerta. Estabelecendo a origem da integral em Eudoxus (390-338 AC) ("*Método de Exaustão*") e, por outro lado, a origem do conceito de derivada nos trabalhos de Newton (1643-1727) e Leibniz (1646-1716), o intervalo de tempo entre os dois eventos soma mais de 2000 anos de Matemática, período em que pontificaram algumas das mentes mais brilhantes da humanidade, como foram as de Arquimedes (287-212 AC) e Galileu (1564-1642), para citar apenas dois extemos. Porque esses e outros luminares da ciência não teriam atinado com o importante e necessário conceito da derivada, que permaneceu dormente até que Newton, em suas próprias palavras, subisse "nos ombros de gigantes" para estabelece-lo definitivamente no corpo da Matemática? O que dizer então dos ombros galgados pelos estudantes contemporâneos ao se matricularem para um primeiro curso de Cálculo? O fato é que o limite singular utilizado para a definição Newtoniana de derivada é muito mais abstrato e difícil de ser apreendido do que o limite de exaustão que tem um objetivo claro como modelo matemático, o cálculo de comprimentos, áreas e volumes, e, portanto, uma razão de ser muito mais intuitiva.

É interessante ressaltar que a avassaladora prevalên-

cia dessa metodologia anti-histórica para o ensino do Cálculo nunca foi totalmente unânime na Matemática. Dentre as vozes de dissensão, encontram-se dois importantes matemáticos do século XX, cujos textos sobre o assunto são preciosidades didáticas, hoje quase desconhecidos: o "*Vorlesungen über Differential- und Integralrechnung*" Springer Verlag, 1924", de Richard Courant, que foi muito influente em alguns poucos centros da elite matemática do século XX, e o "*Entwicklung der Infinitesimalrechnung*", Springer 1946, de Otto Toeplitz, um pequeno livretinho postumamente publicado e injustamente esquecido. (É de se admirar e louvar o discernimento da defunta Editora Globo de Porto Alegre pela publicação em 1951 da tradução ao português do livro de Courant, entitulado "*Calculo Diferencial e Integral*", 2 volumes, um dos raríssimos textos desde então disponíveis a estudantes brasileiros que pretendam iniciar sua formação matemática com solidez. O texto de Toeplitz foi traduzido para o inglês como "*The Calculus: A Genetical Approach*" em 1963. Ambos tem sua origem na escola matemática de David Hilbert em Goettingen e não sub-estimam o interesse do leitor. A julgar por um depoimento de R.Courant, a ideia dessa metodologia remonta a uma sugestão de Leonhard Euler que todavia não a empregou em seu próprio "*Introductio Analysin Infinitesimorum*", 1777).

O argumento de precedência histórica, ainda que contundente, não é o único e nem necessário para nos convencer de que o conceito de derivada é um dos mais difíceis no ensino do Cálculo. Para motivar o aluno ao estudo do estranho limite singular ($\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$), ironicamente, mesmo os textos de Matemática dita "pura", lançam mão de uma vaga ideia de "velocidade instantânea" sem que se apresente sequer de uma razão plausível para que essa quantidade tenha algum interesse. A ubíqua ilustração geométrica da secante que se aproxima da posição de uma tangente, embora mais visível e sugestiva, por outro lado, deixa de ser explorada em sua conexão com o difícil conceito de limite singular. Essas duas importantes motivações têm a sua origem nos trabalhos de Newton e dependem, para tanto, de algum entendimento sobre o conceito de modelo matemático, tema esse que como já observamos, raramente é tratado com o devido cuidado em disciplinas de Matemática, para prejuízo de todos.

Gottfried W. Leibniz (1646-1716), um outro gigante da Matemática, certamente não teve maiores dificuldades para compreender o sentido e o intuito do conceito Newtoniano de derivação, definido como um limite singular de "fluxons". Entretanto, por motivos históricos e filosóficos que valem a pena especular, ele escolheu um caminho completamente diferente para abordar o mesmo assunto. Esta divergência entre Leibniz e Newton na concepção do Cálculo tinha extensões políticas e pessoais que perduraram por toda a vida comum de ambos com repercussões duradouras entre a escola matemática inglesa e a alemã, denominada pelo historia-

dor Ruppert Hall "*guerra entre filósofos*". (Ruppert Hall [1980]).

O método Leibniziano para a apresentação do conceito de derivada, embora produza resultados equivalentes ao método Newtoniano, tem uma motivação e um ponto de partida distintos daquele. Leibniz baseou o seu argumento no conceito de "aproximação assintótica" de funções, por intermédio de funções consideradas como "padrões simples". Sob este ponto de vista unificado e, alguns diriam, simplificado também, ele pôde sintetizar as principais ideias e métodos do Cálculo (limite, continuidade e derivada) e, talvez de maneira presciente, preparar o caminho para generalizá-los a situações muito mais abrangentes do que as funções de variável real. Entretanto, para formalizar suas ideias, Leibniz introduziu o conceito de "infinitesimal" que, à semelhança de vários outros conceitos surgidos no início do Cálculo, era tão misterioso quanto os "fluxons" de Newton. Esse fato é uma vitória parcial do campo Newtoniano na luta intelectual pela conquista da "alma" do Cálculo e das mentes de seus praticantes fez com que o ponto de vista Leibniziano fosse quase varrido dos textos e da teoria do Cálculo até há pouco tempo (Ruppert Hall [1980], Kline[1974], Grabiner[1981]). O que faz a abordagem de Leibniz interessante nos dias de hoje é a possibilidade de dispensar completamente a aura do mistério "infinitesimal" e, ao mesmo tempo, aproveitar suas enormes vantagens intuitivas e de generalidade. Os substitutos contemporâneos para os infinitesimais de Leibniz são os conceitos de "funções escalas" e "ordem de aproximação" que fazem um papel de "medidas" (padrões) para avaliar a "velocidade" de aproximação. (A Análise Não-Standard, desenvolvida por Abraham Robinson na década de 1960, estabelece um fundamento rigoroso para os "infinitesimais" integrando-os ao conjunto de números reais, mas é uma teoria pouco natural, "estática" e que faz uso de intrincadas ferramentas lógicas, nada intuitivas, eliminando assim a maior vantagem do Método de Leibniz. Utilizaremos o caminho mais "cinemático" e intuitivo das funções escalas.)

A origem psicológica da dificuldade na compreensão do conceito de limite de uma função *à la Cauchy* tem, pelo menos, duas razões óbvias e fortes:

1) A sua formulação "*estática*", enquanto a intuição apreende melhor o conceito de limite como um processo essencialmente "*cinético*" e,

2) O fato de que os valores da função podem não se aproximar monotonicamente do valor limite tal como a variável.

É fácil entendermos intuitivamente que funções como, $\frac{-1}{\log x}$, x^p , ($p > 0$), $\frac{1}{\exp(1/x)}$ se aproximem de zero quando a variável x também se aproxima de zero, $x \rightarrow 0$. Isso, porque elas são monotonicamente decrescentes e é possível verificar, com facilidade, que os seus valores

se tornam pequenos (basta calcular um x para o qual $\varphi(x) < \frac{1}{n}$, n inteiro) e progressivamente menores.

As funções monotonicamente decrescentes dessa classe serão denominadas *funções de escala* ou, simplesmente *escalas*, e serão tomadas como *padrões* representativos do conceito de limite, ao mesmo tempo que estabelecerão "*medidas da velocidade de aproximação*" ou "*padrões de convergência*".

As funções de escala terão um papel conceitual em tudo semelhante aos infinitesimais de Leibniz.

Naturalmente, as funções de escala podem assumir expressões muito complexas mas, para o estudo do Cálculo, as escalas definidas por funções elementares (logarítmos, potências, exponenciais) são mais do que suficientes, e mesmo para os especialistas em Análise Matemática. Portanto, a estratégia do Método de Leibniz é utilizar "padrões simples de medida de convergência" para estimar processos mais complicados. Isso torna o conceito de limite Leibniziano mais concreto, pois "mensurável", ao contrário do procedimento lógico e qualitativo de Cauchy. Além disso, como veremos, não há qualquer *perda de generalidade* que pudesse escandalizar os mais ortodoxos! Na verdade, o Método de Leibniz é equivalente ao Método de Cauchy, *apenas* se apresenta sob um ponto de vista muito mais natural e intuitivo.

É interessante observar que o texto de R.Courant mencionado acima inicia o tratamento de limite com uma análise aritmética explícita das convergências de seqüências numéricas elementares (potências, exponenciais e logarítmos), que representarão posteriormente o papel de *estimativas de convergência*, na mesma linha de argumento de Leibniz. O argumento de Cauchy é apresentado páginas adiante. Analogamente, embora a derivada seja apresentada no texto de Courant inicialmente como limite (singular) de quocientes, o ponto de vista de Leibniz é também descrito e amplamente ilustrado logo em seguida, procedimento esse que é repetido no segundo volume para o tratamento de funções de várias variáveis.

O argumento de Leibniz representado por intermédio de funções de escala foi posteriormente generalizado para o conceito de "*aproximação*" ou "*convergência assintótica*" que, apesar de sua importância como Método em várias áreas, especialmente nas aplicações, é relativamente pouco difundido no ensino formal da Matemática, tanto na dita "pura" quanto na desdita.

Uma manifestação recente e entusiasmada pelo resgate do ponto de vista Leibniziano foi, isoladamente, apresentada pelo eminente matemático Donald Knuth, em carta à Sociedade Americana de Matemática (Knuth[1998]), por ocasião de um movimento para a revisão no ensino de Cálculo nos EUA. Em verdade, ainda que esparsamente, essa abordagem vem sendo empregada ao longo da história, mesmo no Brasil. (Ferreira[1990]).

2 Limite e Continuidade *apud* Leibniz: Aproximação Assintótica

Iniciaremos o argumento de Leibniz estudando a questão da *melhor aproximação local* de uma função geral em termos de funções constantes, que é a classe de funções mais simples possível, e que possibilitará redefinir o conceito de limite e continuidade segundo um novo ponto de vista. A questão que se propõe é simples e pode ser descrita da seguinte maneira:

"Qual a função **constante** que **melhor aproxima** os valores $f(x)$, quando x varia em uma vizinhança (local) de x_0 ?"

O conceito de "*melhor aproximação*", segundo Leibniz, não se refere a valores numéricos para os erros "estáticos", mas a uma "*medida de velocidade de convergência*", que sugere uma conotação cinética muito intuitiva associada também ao conceito de *aproximação assintótica*.

Na interpretação geométrica do Cálculo elementar de funções reais o que se pretende é simplesmente determinar uma reta horizontal que melhor aproxime o gráfico de $f(x)$ nas imediações $x_0 + h$ de um ponto x_0 .

É óbvio que todas as funções constantes, bem ou mal, "aproximam" os valores da função $f(x)$, a questão é escolher a "melhor" dentre elas. No presente caso, se c for o valor de uma função constante a ser testada, a qualidade analítica dessa aproximação será avaliada pelo comportamento da função erro: $\varepsilon_0(h) = f(x_0 + h) - c$. O melhor que se pode esperar nesse contexto geral seria que o erro se aproxime de zero quando $h \rightarrow 0$. Assim, os conceitos de melhor aproximação e de limite tornam-se intercambiáveis e o assunto pode ser inicialmente abordado por qualquer um dos dois lados.

Para apresentarmos o assunto sob o ponto de vista de Leibniz substituiremos os seus infinitesimais pelo conceito fundamental de *escalas e ordem*, o que justifica um certo cuidado formal para a sua apresentação.

Definições: Funções de Escalas & Ordem

1- Uma função $\varphi : R^{++} = \{x \in R ; x > 0\} \rightarrow R^{++}$, *monotonicamente decrescente* (isto é, $\varphi(x) < \varphi(y)$ se $0 < x < y$) tal que para qualquer $N > 0$ exista algum valor x , a partir do qual, tenhamos $\varphi(y) < \frac{1}{N}$, para $y < x$,) é denominada "*escala de aproximação*".

(É claro que estamos introduzindo aqui uma classe de funções que representam da forma mais clara possível o conceito intuitivo de $\lim_{s \rightarrow 0} \varphi(s) = 0$ e, a partir das quais é que desenvolveremos o conceito geral de limite abrangendo todos os casos e situações muito mais complicadas).

2- Diz-se que uma escala α é de "*ordem menor do que β* ", (o que se denota das seguintes ma-

neiras: $\alpha \prec \beta$, ou $\alpha = o(\beta)$) se existir uma escala φ tal que $\alpha(s) \leq \varphi(s)\beta(s)$. (É claro que estamos nos referindo, em outros termos, ao fato: $\lim_{s \rightarrow 0} \frac{\alpha(s)}{\beta(s)} = 0$).

Esta ordenação (parcial e estrita) é também interessante que seja estabelecida entre funções gerais f e g , não necessariamente escalas, mas definidas na vizinhança da origem. Para isto tomemos um contexto bem mais geral que amplia a sua utilidade e não interfere na intuição:

3- Diz-se que $f : E \rightarrow F$, (E e F , dois espaços vetoriais normados) é de "ordem menor" do que $g : E \rightarrow F$ nas vizinhanças da origem, e denota-se por, $f = o(g)$ ou por $f \prec g$, se existir uma escala φ tal que $\|f(x)\| \leq \varphi(\|x\|) \|g(x)\|$. (Hardy[1910], de Bruijn [1958]).

Observações: 1- A escala principal que servirá de padrão básico para o Cálculo é a própria variável, ou seja, a função identidade, $\varphi(s) = s$. Exemplos comuns de escalas menos ou mais "rápidas" do que a identidade são as funções: $\psi(s) = |s \log s|$, $\psi(s) = s^p$ ($0 < p$), $\psi(s) = \exp(\frac{-1}{s})$, assim como os seus produtos e suas composições. É fácil ver que $s^p = o(s)$ se $p > 1$ e $s = o(s^p)$ se $p < 1$, assim como $\exp(\frac{-1}{s}) = o(s^p)$, $\forall p > 0$.

2- As funções escala não precisam ser necessariamente contínuas e nem elementares, basta que satisfaçam a condição 1) acima.

3- A relação " \prec " entre escalas é de fato uma ordem parcial, ou seja, é transitiva, mas não é completa. Para verificar isto, considere a função identidade $\varphi(s) = s$ e uma função α definida em intervalos, $[\frac{1}{n+1}, \frac{1}{n})$ alternadamente, pelas expressões $s^{\frac{1}{2}}$ e s^2 .

4- É importante ressaltar que, sendo as funções potência $\varphi_k(s) = s^k$ (k inteiro e positivo) as escalas mais simples e utilizadas, é comum em alguns textos estabelecer a ordem dessas escalas conforme as suas potências, o que confere um sentido exatamente oposto ao empregado acima. Assim, teríamos s^5 de ordem superior a s^2 , quando na verdade, pela definição acima, como $s^5 = (s^3)s^2$, teremos $s^5 \prec s^2$. A definição utilizada aqui tem a vantagem de não se ater ao caso específico das funções potências e ser mais próxima da desigualdade numérica correspondente.

5- Os conceitos de escalas e de ordenação são os fundamentos da Análise Assintótica onde também define-se o conceito de ordem parcial não estrita \preceq :

Definição:

Diz-se que a escala α é no máximo da ordem da escala β e escreve-se $\alpha \preceq \beta$, se existir uma

constante $C > 0$ tal que $\alpha(h) \leq C\beta(h)$ no respectivo domínio. Se também tivermos $\beta \preceq \alpha$ dizemos que α e β têm a mesma ordem e escreve-se $\alpha \approx \beta$.

Diz-se que uma função f definida nas vizinhanças do zero de um espaço vetorial normado é de ordem menor do que h , quando $h \rightarrow 0$, (e a denotamos genericamente por $o(h)$) se existir uma escala φ tal que: $\|f(h)\| \leq \|h\| \cdot \varphi(\|h\|)$. Esta é uma notação proveniente da Análise Assintótica. (Hardy[1910], de Bruijn[1958]).

6- Observe que a soma e o produto de duas funções $o(h)$ é uma função $o(h)$ e que o produto de duas funções $o(h)$ é uma função $o(h^2)$.

É importante sempre lembrar que o símbolo $o(h)$ designa, genericamente, uma função qualquer com a propriedade assintótica descrita acima. Com esse entendimento, tem sentido "rigoroso" a substituição em uma equação um termo $o(h) + o(h)$ por $o(h)$ e outras operações que não teriam sentido no contexto usual.

Definição de Limite e Continuidade segundo Leibniz:

Considere uma função $f : E \rightarrow F$, onde E, F são espaços vetoriais normados. Diz-se que esta função tem limite y_0 para $x \rightarrow x_0$ se em alguma vizinhança de x_0 o erro $\varepsilon(h)$ da aproximação

$$f(x_0 + h) = y_0 + \varepsilon(h),$$

admite uma estimativa $\|\varepsilon(h)\| \leq \varphi(\|h\|)$ para alguma (qualquer !) escala φ e, claro, f é dita contínua em x_0 se $y_0 = f(x_0)$.

Observações: 0- Em termos do argumento de Leibniz, uma função é contínua em um ponto x_0 se admite a melhor aproximação por função constante nas vizinhanças desse ponto e a constante é exatamente $f(x_0)$.

1- Essa definição é totalmente equivalente ao usual critério $\epsilon - \delta$; a definição de Cauchy essencialmente requer a construção implícita de uma função escala $\delta(\epsilon)$, que define os valores de δ em sucessivos intervalos de ϵ , por exemplo, $[\frac{1}{n+1}, \frac{1}{n})$. A vantagem do ponto de vista Leibniziano é sua interpretação geométrica clara e a ênfase no papel da velocidade de aproximação avaliada por alguma escala.

2- É obvio que a escala φ não é única pois na definição não se requer que ela seja a "melhor" no sentido da ordem parcial, (que nem precisa existir), mas simplesmente que seja alguma escala que garanta a convergência! Por exemplo, dentre muitas, a escala $\varphi_C(h) =$

$C\varphi(h)$ (para $C > 1$) e qualquer outra escala Φ de ordem maior do que φ , isto é, $\varphi \prec \Phi$, também cumpre o mesmo papel.

3- Fica claro da definição também que os conceitos de limite e continuidade estão intimamente ligados às normas escolhidas para cada espaço vetorial, E e F , e uma função que é contínua com uma escolha de normas, pode ser descontínua com outra escolha de normas. Todavia, se os espaços E , e F são de dimensão finita, a continuidade da função independe das normas escolhidas, já que são todas equivalentes.

4- O conceito de continuidade segundo Leibniz pode ser definido analogamente para funções em espaços métricos, E e F , que não foram utilizados na definição para evitar a notação mais carregada de distância e pelo fato de que o conceito de derivada a ser tratado em seguida exigirá o contexto de espaços vetoriais. De qualquer forma, o mais importante é a ideia e não o contexto.

3 Derivada de Leibniz-Fréchet

O conceito de derivada Newtoniana, como limite (singular) do quociente entre a variação da função e a variação da variável, perde totalmente o seu significado nos contextos em que não existe a recíproca da variável, como é o caso da maioria das situações, com exceção do Cálculo de uma variável real ou complexa. Os exemplos mais notórios desta dificuldade elementar são os espaços euclidianos R^n , $n > 2$, os espaços de matrizes (ou de operadores lineares) e os espaços vetoriais em geral. (A possibilidade da existência de álgebras com divisão em R^n , para $n > 2$ é uma questão antiga que tem, de certa forma, uma resposta negativa apesar do descobrimento dos quaternions, octonions e etc. v. Birkhoff-McLane[1941]).

A derivada segundo Leibniz, que para funções de uma variável real produz resultados exatamente equivalentes aos da derivada Newtoniana, será facilmente estendida para situações mais abrangentes, como veremos a seguir.

Para uma maior abrangência sem perda de naturalidade, trataremos aqui de funções definidas em espaços vetoriais que podemos entender com sendo R^n .

Uma vez definido o conceito de *melhor aproximação local* por intermédio da classe de funções constantes, o próximo passo natural, segundo o argumento de Leibniz, será utilizar um procedimento análogo considerando-se uma classe expandida de funções aproximantes que admite também as funções lineares (as mais simples depois das constantes) com a finalidade de obter, claro, uma aproximação (assintótica) aperfeiçoada.

Na interpretação geométrica do Cálculo elementar de funções reais o que se pretende agora é determinar uma reta que melhor aproxime os valores da função

contínua $f(x)$ nas imediações $x_0 + h$ de um ponto x_0 . Experimentando geometricamente com uma curva "suave" para gráfico de $f(x)$ em vizinhança de um ponto dado x_0 , concluímos que uma reta genérica que passa pelo ponto $(x_0, f(x_0))$ produz um erro ε_a , $f(x_0 + h) = f(x_0) + ah + \varepsilon_a(h)$, que meramente se aproxima de zero (se f for contínua), o que, sob o ponto de vista "cinético" (*assintótico*), não há aperfeiçoamento algum com relação à aproximação pela reta horizontal.

Entretanto, no "melhor ajuste geométrico" dado pela reta tangente (quando ela existe) observa-se que o erro ε_a se aproxima mais rapidamente de zero do que o desvio h , ou seja, $\varepsilon_a(h) = o(h)$ e, portanto, podemos escrevê-la na forma $\varepsilon_a(h) = h\varepsilon_0(h)$ onde ε_0 se aproxima de zero segundo alguma escala. Esta afirmação é, de fato, um aperfeiçoamento com respeito à aproximação por função constante, pois naquele caso bastava uma estimativa de erro por *qualquer* escala. Se este a for encontrado, verificamos que ele é único, já que podemos escrever:

$$\left| \frac{f(x_0+h) - f(x_0)}{h} - a \right| \leq \varphi(h), \text{ onde } \varphi(h) \text{ é uma escala (qualquer) e isto significa } \lim_{h \rightarrow 0} \frac{f(x_0+h) - f(x_0)}{h} = a, \text{ isto é, a derivada Newtoniana.}$$

Por outro lado, se este último limite existir (isto é, existindo a derivada Newtoniana) então podemos voltar atrás e re-escrevê-la na forma Leibniziana, o que mostra claramente a equivalência das duas definições para funções de variável real.

Passemos então à definição formal de derivada Leibniziana em um contexto mais geral, que inclua situações onde é impossível definir o limite Newtoniano por inexistência do recíproco de h . O nome de Maurice R. Fréchet (1878-1973) é frequentemente associado a essa definição de derivada (com justiça mas sempre sem o de G.W. Leibniz (1646-1716)) por ter sido ele o responsável pela sua (re)descoberta em um contexto abstrato mais de 300 anos após Leibniz.

Definição de Derivada de Leibniz-Fréchet:

Seja $f : E \rightarrow F$ uma função definida entre espaços vetoriais normados E , e F . Se para o ponto x_0 existir uma função *linear contínua* $L_0 : E \rightarrow F$, ($L \in \mathcal{L}(E, F) = \{ \text{Funções lineares contínuas } L : E \rightarrow F \}$), tal que $f(x_0) + L_0(x - x_0)$ aproxima os valores de $f(x_0 + h)$ com um erro de ordem menor do que $\|h\|$,

$$f(x_0 + h) = f(x_0) + L_0h + o(\|h\|),$$

então L_0 é chamada derivada de f segundo Leibniz-Fréchet no ponto x_0 e denotada com um dos símbolos:

$$L_0 = Df(x_0) = \partial f(x_0) = f'(x_0) = \nabla f(x_0) = \frac{\partial f}{\partial x}(x_0).$$

Se f tem derivada de Leibniz-Fréchet em todos os pontos de um conjunto $U \subset E$, então dizemos que f é diferenciável (segundo Leibniz-Fréchet) em U .

Observações: 0- Seguindo a literatura por vezes utilizaremos apenas o termo "Derivada de Fréchet".

1- A existência da derivada de Leibniz-Fréchet inclui implicitamente a existência da aproximação ótima por função constante pois, exige-se que a função linear L_0 seja contínua, isto é, limitada no sentido $\|L_0h\| \leq C \|h\|$. Assim, a expressão $f(x_0) + L_0h$ pode ser considerada como um aperfeiçoamento da aproximação (assintótica) realizada apenas por $f(x_0)$.

2- A derivada de Leibniz-Fréchet $\partial f(x)$, se existir, é única pois, caso contrário, teríamos $L_1(sh) - L_2(sh) = o(\|sh\|)$, de onde, dividindo por s no limite nulo, obtemos $L_1(h) - L_2(h) = 0$ para todo vetor $h \in E$.

3- Se f for diferenciável em $U \subset E$, então a derivada ∂f é uma nova função da forma:

$\partial f: U \rightarrow L(E, F) = \{ \text{Conjunto das funções lineares contínuas } L: E \rightarrow F \}$, que por sua vez pode ser também derivada se o espaço vetorial $L(E, F)$ for dotado de uma norma, por exemplo, a uniforme. Assim, podemos definir, ainda que progressivamente de maneira mais complexa, derivadas de qualquer ordem.

4- No caso de funções $f: U \subset R^n \rightarrow R^m$ a derivada de Leibniz-Fréchet em cada ponto, $\partial f(x)$, é uma transformação linear entre R^n e R^m , e, portanto, identificável com uma matriz $M_{n \times m}(R)$ que a representa (chamada matriz "Jacobiana"), de onde podemos escrever: $\partial f: U \subset R^n \rightarrow M_{n \times m}(R)$. Observe que $M_{n \times m}(R)$ é um espaço vetorial de dimensão n^2 que pode ser normado de diversas formas (equivalentes), a mais comum sendo a norma euclidiana (Frobenius): $\|A\|_2 = \sqrt{\text{Tr}(AA^t)} = \sqrt{\sum A_{kj}^2}$.

5- Se existir a derivada de Leibniz-Fréchet, então é fácil ver que existem todas as derivadas direcionais (Gateaux) ∂_v e, $\partial_v f(x_0) = \frac{\partial f}{\partial x}(x_0)v$. Em particular, se $f: R^n \rightarrow F$, então existem todas as derivadas parciais, e $\frac{\partial f}{\partial x_k}(x_0) = \frac{\partial f}{\partial x}(x_0)e_k$. A recíproca entretanto não é verdadeira pois pode não haver uniformidade (com relação à direção) para os limites que definem as derivadas direcionais. Mais adiante mostraremos um importante Teorema que fornece condições úteis nas derivadas direcionais para que exista a derivada de Leibniz-Fréchet.

6- Se $f: R^n \rightarrow R^m$, então a matriz que representa a função linear $\partial f(x): R^n \rightarrow R^m$, pode ser calculada

via derivadas direcionais e verificamos facilmente que $[\partial f(x)]_{ij} = \frac{\partial f_i}{\partial x_j}(x)$, a chamada Matriz Jacobiana de f no ponto x .

7- Por outro lado, se existir uma função L não contínua que satisfaz à condição de aproximação acima estabelecida, não teremos a derivada de Leibniz-Fréchet, mas teremos, sim, todas as derivadas direcionais (de Gateaux).

8- Se $f: U \subset R^n \rightarrow R$ for uma função diferenciável segundo Leibniz-Fréchet, então, $\partial f: U \subset R^n \rightarrow L(R^n, R) \approx M_{1 \times n}(R)$, ou seja, $\partial f(x)$ pode ser representada por uma matriz $1 \times n$, $[\partial f(x)]_k = \frac{\partial f}{\partial x_k}(x)$. Mas, segundo o Teorema de Riesz-Fréchet (o mesmo Fréchet) também podemos representar o produto matricial Lh como um produto escalar, $Lh = \langle l, h \rangle, \forall h \in R^n$, onde $l \in R^n$, claro, é dado por $l_k = L_{1k}$. Com esta representação, podemos interpretar $\partial f(x)$ como sendo um vetor que chamaremos de "Gradiente de f em x ". Frequentemente os símbolos $\partial f(x) = \nabla f(x) = \text{grad} f(x)$, são utilizados para enfatizar este aspecto, mas não há uma concordância geral para estas notações.

9- Em geral, se $f: U \subset H \rightarrow R$, for uma função definida em um espaço de Hilbert H real, e com valores reais, então, como ainda vale o Teorema de Representação de Riesz-Fréchet neste contexto, podemos identificar a derivada de Fréchet $\partial f(x)$ como sendo um vetor de H , ou seja, $\partial f: U \subset H \rightarrow H$, e $f(x+h) = f(x) + \langle \partial f(x), h \rangle + o(h)$. Este teorema continua valendo para funções escalares definidas em Espaços de Hilbert complexos: $f: U \subset H \rightarrow C$. A representação gradiente de Riesz-Fréchet para a derivada de uma função escalar definida em um espaço de Hilbert será de grande utilidade no tratamento de funcionais e fundamental para a interpretação de vários resultados em EDPs.

10- O conceito de Aproximação Assintótica é uma extensão e uma particularização da ideia de aproximação introduzida por Leibniz. Se dispomos de uma sequência de funções (não necessariamente escalas) $\{\varphi_k\}$ definidas em uma vizinhança da origem e ordenadas na forma $\varphi_{k+1} \prec \varphi_k$, então dizemos que a função f (também definida nesta vizinhança) tem uma expansão assintótica nesta classe se existirem coeficientes $\{c_k\}$ e

$$\left| f(x) - \sum_{k=0}^{k=N} c_k \varphi_k(x) \right| \prec \varphi_N, \text{ e, neste caso, escrevemos, } f(x) \approx \sum_{k=0}^{\infty} c_k \varphi_k(x). \text{ (de Bruijn[1958]).}$$

4 Exemplos de Derivadas de LEIBNIZ-FRÉCHET

A explicitação de derivadas de Fréchet para funções vetoriais relevantes sempre produz resultados que por si

mesmos já se constituem em Teoremas notáveis da Matemática e importantes para suas aplicações. Nessa seção obteremos algumas dessas derivadas diretamente a partir da caracterização Leibniz-Fréchet.

4.1 Funções Lineares, Quadráticas e Multilineares (Polinomiais)

Esses exemplos, de tão simples, serão apenas citados e o leitor poderá verificar a veracidade deles sem dificuldade e, com proveito.

a- A derivada de uma função Linear $L : E \rightarrow F$ contínua é constante e igual a própria L , ou seja, $\partial L = L$.

(A aparente incongruência entre igualdade acima e a tradicional $\frac{d e^x}{dx} = e^x$ deve ser esclarecida pela correta interpretação de cada caso. Observe que no Cálculo elementar a derivada é representada por um número que descreve a inclinação da reta tangente e no caso exponencial é variável com o ponto x , ao contrário da derivada da função linear.)

b- $E = H$ um espaço de Hilbert, S uma operação linear auto-adjunta contínua, e a forma quadrática $Q(x) = \langle Sx, x \rangle$, $Q : E \rightarrow C$. Então,

$\partial Q(x)h = \frac{1}{2} \langle Sx, h \rangle$, que, na representação de Riesz-Fréchet, pode ser re-escrita como gradiente na forma: $\partial Q(x) = \nabla Q(x) = \frac{1}{2} Sx$.

c- $B : E \times E \rightarrow F$, é função bilinear contínua, ou seja, B é linear em cada variável e existe $C > 0$ tal que, $\|B(x,y)\| \leq C \|x\| \|y\|$. A derivada de $b : E \rightarrow F$, $b(x) = B(x,x)$ pode ser facilmente calculada assim como $\partial B(x,y)$. Exemplo elementar mas interessante destas é o produto vetorial em R^3 .

d- O produto vetorial misto pode ser encarado como uma como uma função trilinear $D : R^3 \times R^3 \times R^3 = M_3(R) \rightarrow R^3$, sua derivada de Fréchet tem interesse especial.

4.2 Derivada da função matricial inversão (recíproca), $r : I_n(C) \rightarrow I_n(C)$, $r(A) = A^{-1}$ - Séries de Neumann:

Consideremos agora o espaço vetorial das matrizes complexas de ordem n que denotaremos por $M_n(C)$ com a norma de Frobenius: $\|A\|_2 = \sum_{k,j} |A_{kj}|^2$, proveniente do produto interno euclideano, $\langle A, B \rangle = Tr \{AB^*\} = \sum_{k,j} A_{kj} \overline{B_{kj}}$, onde Tr é o traço e o subconjunto das matrizes k,j inversíveis, que denotaremos por $I_n(C)$. Se $f : C \rightarrow C$ for uma função analítica qualquer definida em uma vizinhança da origem, ou seja, representável (segundo Weierstrass) por uma série de potências $f(z) = \sum_{k=0}^{\infty} a_k z^k$, convergente para $|z| < R$, então, podemos definir as

funções matriciais $f(A)$ desde que $\|A\|_2 < R$, como limite das séries de potências: $f(A) = \sum_{k=0}^{\infty} a_k A^k$. Para tanto, basta utilizar a desigualdade $\|A^p\|_2 \leq (\|A\|_2)^p$, $|A_{kj}| \leq \|A\|_2$, e o critério de majoração de Weierstrass. (Bassanezi-Ferreira[1988]).

Em particular, utilizando a expansão em série geométrica da função $f(z) = \frac{1}{1-z} = \sum_{k=0}^{\infty} z^k$, (para $|z| < 1$), podemos definir a função recíproca matricial r nas imediações de uma função $A \in I_n(C)$ pela famosa série de Carl Neumann:

$$\begin{aligned} r(A+H) &= (A+H)^{-1} = ((I+HA^{-1})A)^{-1} \\ &= A^{-1} (I+HA^{-1})^{-1} = A^{-1} \sum_{k=0}^{\infty} (-HA^{-1})^k \\ &= A^{-1} - A^{-1}HA^{-1} + A^{-1}(HA^{-1})(HA^{-1}) + \dots \\ &= A^{-1} - A^{-1}HA^{-1} + o(H), \end{aligned}$$

desde que $\|HA^{-1}\| < 1$, o que é satisfeito se $\|H\| < \frac{1}{\|A^{-1}\|}$, ou seja, neste caso, $(A+H) \in I_n(C)$.

A própria expansão de Neumann na forma $(A+H)^{-1} = A^{-1} - A^{-1}HA^{-1} + o(H)$ fornece explicitamente a derivada e Leibniz-Fréchet da função r no valor A , ou seja,

$$\partial r(A)H = -A^{-1}HA^{-1}.$$

É interessante observar que esta expressão guarda uma certa semelhança com a expressão para a derivada da função recíproca $r(z) = \frac{1}{z}$, real ou complexa, pois, $r'(a) = \frac{-1}{a^2} = -(a^{-1})^2$. A ressalva é de que para $n > 1$ as matrizes não são, em geral, comutativas e, portanto, devem guardar suas respectivas ordens nos produtos.

Do resultado acima conclui-se que as funções matriciais analíticas $f(A)$ não tem, necessariamente, derivadas de Leibniz-Fréchet dadas por $f'(A)$ como poderia ser ingenuamente previsto. Na verdade, essas derivadas raramente podem dispor de representações elementares devido em parte à não comutatividade do produto matricial. Esse e os próximos exemplos são casos excepcionais. Outros casos mais simples que podem ser verificados diretamente pelo leitor são as funções de potências inteiras: $f(A) = A^n$.

4.3 Derivada da função determinante $\det : I_n(C) \rightarrow C$ - Fórmula de Euler-Schwinger

Consideremos a função determinante com variável matricial e valores reais definidas no subconjunto das matrizes complexas inversíveis, $n \times n$, $\det : I_n(C) \rightarrow C$.

Inicialmente observemos que a derivada de Fréchet da função \det em cada ponto $A \in I_n(C)$ é uma função

linear de valores complexos definida no espaço das matrizes $M_n(\mathbb{C})$ e \mathbb{C} , e, portanto, pelo teorema de representação de Riesz-Fréchet pode ser escrito na forma:

$$\partial (\det) (A)H = \text{Tr}(\Lambda H^*) = \langle \Lambda, H \rangle,$$

onde, naturalmente, a matriz Λ , em geral, depende de A .

Para determinarmos a derivada de Fréchet, analisemos a expansão:

$$\begin{aligned} \det(A + H) &= \det\left((I + HA^{-1})A\right) \\ &= \det A \det\left(I + HA^{-1}\right), \end{aligned}$$

e nos fixemos no termo $\det(I + HA^{-1}) = \det(I + \Omega)$, lembrando que todos os termos da matriz $\Omega = HA^{-1}$ contêm fatores de termos H_{kj} da matriz H . Estamos interessados em expandir a (entranhada) expressão $\det(I + \Omega)$, mas, felizmente, apenas nos seus termos que não contenham H_{kj} 's, ou que tenham H_{kj} 's apenas em primeiro grau, pois $|H_{kj}^2| \leq \|H\|^2, |H_{kj}^3| \leq \|H\|^3$ etc. Lembrando que o determinante de uma matriz é calculado como a soma de parcelas, cada uma delas, sendo um produto de n fatores onde aparecem representantes de todas as linhas e todas as colunas, concluímos que se em um fator houver mais de um termo do tipo Ω_{kj} este termo será $o(\|H\|)$. Assim, a única parcela que pode nos interessar é aquela formada pelo produto dos termos da diagonal de $I + \Omega$, isto é: $(1 + \Omega_{11}) \cdots (1 + \Omega_{kk}) \cdots (1 + \Omega_{nn}) = 1 + \text{Tr}(\Omega) + o(\|H\|)$.

Voltando à expressão original, escrevemos:

$$\begin{aligned} \det(A + H) &= \det A \det\left(I + HA^{-1}\right) \\ &= \det A \left(1 + \left(\text{Tr}(HA^{-1})\right)\right) + o(\|H\|) \\ &= \det A + \det A \left(\text{Tr}(HA^{-1})\right) + o(\|H\|), \end{aligned}$$

de onde tiramos a importante fórmula (obtida por Leonhard Euler no século XVIII para a obtenção das equações da dinâmica de fluidos que lhe deu eterna fama, e "redescoberta" por Julian Schwinger no século XX no seu trabalho sobre as equações da eletrodinâmica quântica que lhe rendeu eterna fama e um terço do prêmio Nobel de 1958, junto com R.Feynman e S.Tomonaga):

$$\partial \det(A)H = \det A \left(\text{Tr}(HA^{-1})\right) = \langle \Lambda, H \rangle,$$

ou, segundo a representação "gradiente" de Riesz-Fréchet

$$\frac{\partial \det}{\partial A}(A) = \text{grad}(\det(A)) = \Lambda(A) = (\det A) \left(A^{-1}\right)^*.$$

Essas fórmulas podem ser estendidas para a Análise Operacional, que é o estudo de funções cujas variáveis são operadores lineares em espaços de Hilbert e operadores diferenciais. (Lax[2000], Nazaikinskii&al[1998], Maslov&al.[1980]).

4.4 Derivada da função exponencial de operadores lineares $\partial(e^A)$: Fórmula de Dyson

A função exponencial é certamente a mais importante função não-aritmética (isto é, não definível por uma série finita de operações aritméticas), uma vez que as funções logarítmica e todas as trigonométricas podem ser definidas a partir dela. A definição de e^z no domínio complexo (seu habitat natural) pode ser realizada de várias maneiras equivalentes que são importantes para enfatizar seus diferentes aspectos, em particular:

$e^z = \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^{\infty} \frac{z^k}{k!} = w(1)$, onde $w(t)$ é solução da equação diferencial ordinária no plano complexo: $\frac{dw}{dt} = zw, w(0) = 1$, além de outras.

As três expressões fazem sentido e são iguais quando $z = A$ for uma matriz quadrada ou mesmo para operadores lineares muito mais gerais, tais como operadores diferenciais, desde que sob condições específicas. Apresentando esse item no contexto das matrizes evita-se o emaranhado técnico dos casos mais gerais, ao mesmo tempo que se expõe todo o formalismo da Análise (ou, Cálculo) Operacional, que é o estudo de funções cujas variáveis são operadores. (Maslov[1980], Lax[2000], Ferreira-Fassoni[2014]). A derivada da função exponencial complexa pode ser obtida tanto pela definição Newtoniana quanto Leibniziana. Entretanto, a derivada desta função no contexto da Análise Operacional somente poderá ser definida pelo Método de Leibniz que será ilustrado a seguir. A sugestão mais natural para isto seria a representação por série. Entretanto, ao escrevermos $e^{A+H} = \sum_{k=0}^{\infty} \frac{(A+H)^k}{k!}$ observa-se que cada potência $(A + H)^k = (A + H) \cdots (A + H)$ produz k termos lineares em H e, pelo menos à primeira vista, seria intratável colecioná-los todos em uma expressão razoável. Enfim, não vale aqui o resultado do Cálculo elementar, talvez ingenuamente esperado, de que a derivada da exponencial é a própria exponencial!

Para obter uma expressão razoável que represente a derivada, utilizaremos a sua caracterização como solução de uma equação diferencial ordinária operacional (Bassanezi-Ferreira[1988]): $\frac{dU}{dt} = AU, U(0) = I$, de onde, $e^A = U(1)$. Assim, $e^{A+H} = V(1)$ onde $V(t)$ é solução da equação: $\frac{dV}{dt} = (A + H)V = AV + HV, V(0) = I$, que pode ser também escrita na forma $V(t) = e^A + \int_0^t e^{(t-s)A}HV(s)ds$ (que se obtém resolvendo a equação modificada com a fatoração operacional $(e^{tA}) \frac{d}{dt} (e^{-tA}) = \left(\frac{d}{dt} - A\right)$ e o teorema fundamental do Cálculo). Substituindo duas vezes essa expressão para $V(t)$ nela mesma, obtemos: $V(t) = e^{tA} + \int_0^t e^{(t-s)A}He^{sA}ds + o(H)$, de onde

o argumento de Leibniz produz a importante Fórmula de Dyson:

$$\partial (e^A) H = \int_0^1 e^{(1-s)A} H e^{sA} ds.$$

Fórmulas para derivadas, particularmente as da Análise Operacional, são importantes no cálculo de aproximações assintóticas. (Hoppensteadt[2000], Trenoguine [1985], Maslov&al.[1980]).

A fórmula de Dyson é surpreendente por relacionar dois aspectos muito distintos da Matemática, o combinatório (de quantas formas o termo linear aparece na expansão em série de potências) com uma expressão contínua. Em geral, sempre que isso acontece em Matemática alguma coisa muito profunda espreita submersa na conexão. (Essa fórmula foi originalmente obtida pelo físico Freeman Dyson em 1948 no seu trabalho de fundamentação da eletrodinâmica quântica, pelo qual ele não compartilhou o prêmio Nobel dado a J.Schwinger, R.Feynman e S.Tomonaga ganho pelo mesmo motivo).

4.5 Decomposição Polar da Derivada de Leibniz-Fréchet e a Derivada Complexa de Cauchy: Equações de Cauchy-Riemann

Uma função $f : R^2 \rightarrow R^2$, definida coordenada a coordenada por: $f(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2))$ pode sempre ser considerada como uma função de variável complexa e valores complexos, escrevendo-se, naturalmente, $z = x_1 + ix_2$ e, $f(z) = f_1(z) + i f_2(z)$, e, vice-versa, tal como fazia Cauchy.

É um fato notório em Matemática que a dimensão $n = 2$ é um exemplo excepcional de espaço vetorial euclideo, R^n , onde se pode definir a recíproca de qualquer vetor $h \neq 0$, utilizando-se para isto a sua estrutura natural de corpo complexo, por intermédio da identificação: $R^2 \approx C, (x_1, x_2) \sim z = x_1 + i x_2$.

Portanto, analogamente a $n = 1$, podemos também neste caso definir a derivada Newtoniana de funções f como um limite (quando existir) da razão entre a diferenças funcional e da variável, a chamada derivada complexa de Cauchy: $f'(z_0) = \lim_{h \rightarrow 0} \frac{f(z_0+h) - f(z_0)}{h}$.

Utilizando aqui o já conhecido argumento do Cálculo real, é possível reescrever a expressão limite que define a derivada de Cauchy no sentido de Leibniz:

$$f(z_0 + h) = f(z_0) + f'(z_0)h + o(h),$$

onde o termo $Lh = f'(z_0)h$ (resultante de um produto complexo) é linear em h . Concluímos então imediatamente que a transformação linear L é também uma derivada de Leibniz-Fréchet da função f em z_0 .

Mas, como se sabe, uma transformação linear em R^2 obtida pelo produto complexo com um número fixo $\alpha = re^{i\theta}$, $Lh = \alpha h = re^{i\theta}h$, pode ser geometricamente interpretada como composição de uma Rotação ($e^{i\theta}$) e uma Homotetia (ou, vice-versa), e se $f'(z_0) = \alpha = a + ib$ e $h = h_1 + ih_2$, este produto complexo pode ser representado matricialmente da seguinte forma:

$$f'(z_0)h = (ah_1 - bh_2) + i(ah_2 + bh_1) \approx \begin{pmatrix} ah_1 - bh_2 \\ ah_2 + bh_1 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}.$$

Por outro lado (o Leibniziano) a derivada de Fréchet para $f(x_1, x_2)$ pode ser representada também pela sua matriz jacobiana

$$\frac{\partial f}{\partial x}(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix},$$

e, como a derivada de Fréchet-Leibniz é única, concluímos que, existindo a derivada de Cauchy, existirá também a derivada de Leibniz-Fréchet e a suas representações complexas e de Jacobi deverão ser iguais, ou seja, necessariamente devemos ter as:

Equações (de Compatibilidade) de Cauchy-Riemann:

$$a = \frac{\partial f_1}{\partial x_1} = \frac{\partial f_2}{\partial x_2},$$

$$b = -\frac{\partial f_1}{\partial x_2} = \frac{\partial f_2}{\partial x_1}.$$

Por outro lado, se a função f tiver derivada de Leibniz-Fréchet no ponto $x_0 = (x_{10}, x_{20})$ e a sua representação matricial Jacobiana for da forma $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, ou seja, se forem válidas as condições de compatibilidade de Cauchy-Riemann, então podemos reescrever a expressão matricial na forma de um produto complexo

$$\begin{pmatrix} \partial_{x_1} f_1(x_0) & \partial_{x_2} f_1(x_0) \\ \partial_{x_1} f_2(x_0) & \partial_{x_2} f_2(x_0) \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \left(\frac{\partial f_1}{\partial x_1} + i \frac{\partial f_2}{\partial x_1} \right) (h_1 + ih_2),$$

e daí concluir que esta função f também admite a derivada de Cauchy-Riemann, de onde vem o importantíssimo Teorema de Cauchy-Riemann.

4.5.1 Teorema de Cauchy-Riemann:

Uma função $f : R^2 \rightarrow R^2$ diferenciável segundo Fréchet-Leibniz é diferenciável segundo Cauchy-Riemann (quando interpretada como função $f : C \rightarrow C$) se e somente se forem válidas as Equações (condições de compatibilidade) de Cauchy-Riemann.

Dessa abordagem podemos facilmente obter equações de compatibilidade de Cauchy-Riemann para qualquer tipo de coordenadas (Ferreira-Borges[2014]):

1- Considere uma função $f : R^2 \rightarrow R^2$ definida com as variáveis em coordenadas polares e valores em coordenadas cartesianas, ou seja, $f(r,\theta) = f_1(r,\theta) + if_2(r,\theta)$. As condições de compatibilidade de Cauchy-Riemann para as funções $f_k(r,\theta)$ podem ser obtidas seguindo o argumento acima.

2- Considere uma função $f : R^2 \rightarrow R^2$ definida com as variáveis e valores em coordenadas polares, ou seja, $f(r,\theta) = R(r,\theta) \exp(i\Theta(r,\theta))$. As condições de compatibilidade de Cauchy-Riemann para as funções $R(r,\theta)$ e $\Theta(r,\theta)$ seguem do argumento acima.

3- Considere que a função $f : R^2 \rightarrow R^2$ e as coordenadas curvilíneas definidas por, $\varphi : R^2 \rightarrow R^2$, e $\psi : R^2 \rightarrow R^2$ (ambas com inversas, pelo menos locais, sendo $\varphi, \varphi^{-1}, \psi, \psi^{-1}$, todas diferenciáveis segundo Fréchet). A expressão generalizada para as condições de compatibilidade de Cauchy-Riemann nestas novas coordenadas também seguem diretamente do argumento acima.

4.6 Decomposição de Cauchy-Helmholz: Operadores Rotacional e Divergente

Uma função definida entre dois espaços R^n pode ser interpretada geometricamente como diversos modelos matemáticos, o que leva a diferentes interpretações também para os vetores x e $f(x) = y$, assim como a distintas motivações no estudo da própria f . Em um exemplo mais comum f é interpretada como mapeamento que descreve pontos y de uma cópia de R^n por intermédio de pontos x de outra cópia do mesmo R^n . Nesse caso, tanto x como y representam *posições* em seus respectivos domínios. (É interessante observar que para esse modelo a função f é bijetora embora o uso do termo "mapeamento" para f em matemática prescindia dessa condição). Em um segundo exemplo, x representa um *ponto* e $f(x) = y$ representa um *deslocamento* deste ponto, ou seja, $u(x) = x + f(x)$ descreve uma *deformação* do espaço, cuja motivação tem a sua origem em teoria da elasticidade.

A interpretação, que será utilizada nessa seção é oriunda também da dinâmica do meio contínuo, mais particularmente de fluidos. Nesse caso, x representará um *ponto* do espaço e $f(x)$ a sua *velocidade*, ou seja, diz-se que f é um campo de *velocidades*. Em seus estudos de dinâmica de fluidos, o já citado matemático A.-L.Cauchy e o físico H.Helmholz empregaram uma decomposição muito simples da derivada de um campo de velocidades que põe em relevo propriedades importantes do movimento do fluido nas imediações de um ponto e esclarece a natureza de algumas operações diferenciais do Cálculo Vetorial, que usualmente são

impostas como receituário com nomes sugestivos mas misteriosos. Apesar da motivação essencialmente física desse procedimento, o resultado obtido é de importância fundamental para a Análise. Assim, consideremos a expressão de Leibniz para o valor de $f(x_0 + h) = f(x_0) + \partial f(x_0)h + o(h) = v_0 + Mh + o(h)$ e concentremos nossa atenção na sua aproximação de primeira ordem, $v_0 + Mh$, que reescreveremos como

$$v_0 + Ah + Sh,$$

em que a matriz $M = \partial f(x_0)$ é descrita pela decomposição, quase óbvia, de Cauchy-Helmholz, $M = \frac{1}{2}(M - M^t) + \frac{1}{2}(M + M^t)$, observando-se que $A = \frac{1}{2}(M - M^t)$ é uma matriz anti-simétrica e $S = \frac{1}{2}(M + M^t)$ é uma matriz simétrica. Se restringirmos agora, por simplicidade, esse exemplo ao espaço físico com dimensão $n = 3$, podemos utilizar um antigo resultado (de Geometria Analítica ou Cinemática Elementar), pouco conhecido hoje, mas, crucial em vários contextos:

Lema de Euler:

"Uma operação linear em R^3 representada em base ortonormal por uma matriz anti-simétrica A pode ser também representada como um produto vetorial, ou seja, existe um vetor ω (único, a menos de orientação do espaço) tal que $Ah = \omega \times h$, e vice-versa".

Continuando com a mesma interpretação geométrica, podemos interpretar, com Euler, o campo de velocidades $Ah = \omega \times h$ como sendo um campo de rotação em torno do eixo ω com velocidade angular $|\omega|$ na orientação escolhida. Resumindo, a decomposição de Cauchy-Helmholz para a aproximação de primeira ordem de um campo de velocidades fornece a seguinte interpretação:

O movimento na vizinhança do ponto x_0 consiste na composição de três movimentos: 1) Uma translação comum a todos os pontos $x = x_0 + h$ dada por $v_0 = f(x_0)$, 2) Uma rotação dessa mesma vizinhança (como a de um corpo rígido) definida pelo vetor ω e, 3) Um campo Sh que, por resultados anteriores pode ser interpretado como gradiente de uma função quadrática $Q(x) = \langle x - x_0, S(x - x_0) \rangle$.

Um pouco de manipulação, que não apresentaremos aqui, mas que deveria constar de todos os textos de Cálculo de várias variáveis, mostra que $\omega = \frac{1}{2} \text{rot} f(x_0)$ e que o terceiro termo é o único que imprime uma deformação local no meio (já que os dois primeiros são movimentos de corpo rígido) cuja taxa específica de compressão volumétrica é dada pelo $\text{div} f(x_0) = \text{tr} M = \sum \frac{\partial f_k}{\partial x_k}(x_0)$.

Teoremas integrais de Green, Kelvin e Stokes e outros podem ser obtidos de forma muito mais natural, com mais substância conceitual e generalizáveis para dimensão superior se os operadores diferenciais *rot* e *div* forem apresentados por intermédio da derivada de

Leibniz-Fréchet (Ferreira[1990], Ferreira-Fassoni[2014]).

Com isso, quebra-se de maneira relativamente simples os mistérios dos operadores "rotacional" e "divergente" tão bem guardados na maioria dos textos de Cálculo Vetorial, onde quase sempre aparecem na forma de meros símbolos ociosos de significado e destinados à aplicação de um receituário mecânico.

5 Conclusões

O conceito de derivada tem sido amplamente generalizado de várias maneiras com a finalidade de estender os bem sucedidos métodos de Cálculo a contextos em que as funções tem um domínio ou um significado mais amplo do que o das funções elementares de variável real. Exemplos dessas extensões são a Derivada Geométrica em variedades diferenciáveis (Dubrovnik&al.[1986]), Derivadas Generalizadas de Friedrichs-Sobolev-Schwartz (Lutzen[1982], Lax-[2002]) e Sato (Graf[2000]), Derivada de Fubini, (Lax-[2002]), Derivada Complexa (Henrici[1974]), Derivada Convexa (Trenoguine [1989]), Derivadas Geométricas (Grenander&al.[2000], Delfour&al.[2001], Henry[1982]), q -derivada (Kac&al.[2002]), Derivada Fracionária (Zaslavsky [1998]), Operadores Pseudo-Diferenciais (Nazaiinskii&al[1998]), além da teoria de Leibniz-Fréchet (Trenoguine[1989], Lax[2002], Ferreira-Fassoni[2014]), enfatizada nos itens anteriores.

Em Cálculo elementar de funções reais de variável real é notória a importância do mais famoso e útil dos Teoremas de Fermat que estabelece critérios para pontos de extremo de uma função diferenciável, ou que caracteriza o crescimento ou decréscimo e a inversão local de uma função em termos da sua derivada. Analogamente, diversos aspectos cruciais do comportamento de uma função f definida em espaços vetoriais podem também ser obtidos analisando-se a derivada de Leibniz-Fréchet $\partial f(x_0)$ segundo o seu papel na aproximação $f(x+h) = f(x) + \partial f(x_0)h + o(h)$. Esse procedimento é a base para a formulação de várias teorias de importância fundamental na Matemática contemporânea com grande relevância para as aplicações. Métodos de Perturbação (Keller[1985], Hoppensteadt[2000], Trenoguine[1985]), Análise de sensibilidade (Caswell [2009]), Análise Matricial (Absil&[2007]) e Análise de Estabilidade (Trefethen&al.[2005]), estão entre eles. Em particular, informações cruciais sobre o comportamento local da função f nas imediações do ponto x_0 são extraídas da decomposição do operador linear $\partial f(x_0)$, tal como a possibilidade de inversão local da função f (quando atuante entre um mesmo espaço), e a análise de diversos problemas de Bifurcação, como nas Teorias de Lyapunov-Schmidt e de Hopf.(Keller[1985], Hoppensteadt[1990], Trenoguine[1985], Ambrosetti[1997]).

Os Teoremas de Cauchy-Riemann e Cauchy-Helmholz foram deduzidos a partir de decomposições apropriadas da derivada de Leibniz-Fréchet. A aplicação do conceito de derivada de Leibniz-Fréchet a funcionais (isto é, operadores que atuam em espaços de funções) é uma técnica importante na teoria de equações diferenciais parciais e teorias Físicas (Cvitanovic[2000], Ferreira-Fassoni[2014], Lax[2000]). A algebrização do conceito de derivada também foi sugerida pelo mesmo Leibniz que, em um de seus trabalhos, tratou a operação funcional derivada como um objeto matemático por si mesma passível de ser ela própria submetida a operações, como a soma, multiplicação por escalar e produto (composição). Esse procedimento é a própria essência da álgebra operacional e antecipa a Análise Operacional por vários séculos. Para tanto, Leibniz despojou a sua carregada, mas explicatória, notação quociente $(\frac{dy}{dx})$ de quatro de seus símbolos e a designou simplesmente por um único símbolo d . Talvez por ser tão inovadora, essa abordagem permaneceu obscura na Matemática e suas aplicações até que os notáveis trabalhos de Oliver Heaviside (1850-1925) de maneira totalmente intuitiva e independente, a empregasse com desenvoltura para o desenvolvimento da Teoria Eletromagnética de J.C. Maxwell ao final do século XIX. O enorme sucesso de seus resultados não foi suficiente para despertar o interesse matemático pela formalização rigorosa de seus métodos e nem mesmo foram adotados com entusiasmo na Matemática Aplicada, o que seria de se esperar, dada a sua simplicidade e eficiência. (A propósito, causa ou efeito disso, Heaviside nunca pertenceu aos meios acadêmicos e foi um autodidata excêntrico que morreu quase na miséria. (Nahin[2002]). O aspecto mais elementar da álgebra operacional pode ser facilmente apresentado de maneira "rigorosa" e sem pedantismos, embora, mesmo assim, seja pouco utilizada em textos de equações diferenciais e de diferenças, em que poderia ter maior impacto pedagógico). (Bassanezi-Ferreira[1988]).

A Álgebra Operacional foi desenvolvida matematicamente por Jan Mikusinski (1913-1987), dentre outros, que a colocou em bases sólidas, e úteis, em meados do século XX (Mikusinski[1953]). O desenvolvimento dos chamados Métodos Operacionais que estendem consideravelmente o escopo da Álgebra Operacional se deu vigorosamente durante o século XX impulsionados em grande parte pela influência de teorias Físicas e das Equações Diferenciais Parciais. (Maslov[1980], Nazaiinskii&[1995], Lax[2000]).

Em um segundo artigo sobre esse tema geral, trataremos de alguns tópicos dentre os mais elementares mencionados acima.

Referências

- Ambrosetti, L. (1997). *A Primer of Non Linear Functional Analysis*. Cambridge Univ. Press.
- Absil, P.A., Mahonny, J., Sepulchre, R. (2007). *Optimization Algorithms on Matrix Manifolds*. Princeton Univ. Press.
- Birkhoff, G., MacLane, S. (1941). *A Survey of Modern Algebra*. McMillan.
- Caswell, H. (2009). Sensitivity and Elasticity of density dependent population models, *J. Diff. Eq. Appl.*, 15, 349–369.
- Civitanovic, P. (2001). Niels Bohr Institute, *Online: <http://www.nbi.dk/ChaosBook>*.
- Courant, R. *Cálculo Diferencial e Integral*, 2 volumes, Ed. Globo, Porto Alegre, 1958.
- Bruijn, N. *Asymptotic Methods in Analysis*, North-Holland, 1958.
- Delfour, M., Zolesio, J. *Shapes and Geometries, Analysis, Differential Calculus and Optimization*, SIAM, 2001.
- Dubrovin, B., Novikov, S.P. *Méthodes de la Géométrie Contemporaine*, MIR 1985.
- Ferreira, W.C. Jr. *Notas de Cálculo I*, (não publicadas)-IMECC-UNICAMP, 1990.
- Ferreira, W.C. Jr. (2014). *As Múltiplas Faces da Derivada II: Decomposição*, (a ser publicado).
- Ferreira, W.C. Jr., Fassoni, A.C. (2014). *Análise Funcional: Princípios, Métodos e Fins*, (a ser publicado).
- Ferreira, W.C. Jr., Borges, F.C.M. (2014). *Análise Complexa: Princípios, Métodos e Fins*, (a ser publicado).
- Gelfand, I.M., & al. (1968). *Generalized Functions*, 5 volumes. Academic Press.
- Grabner, J.V. (1981). *The Origins of Cauchy's rigorous Calculus*. MIT Press.
- Graf, U. (2000). *Hyperfunctions: An Introduction*. Birkhauser.
- Grenander, U., Miller, M. (2007). *Pattern Theory: From Representation to Inference*. Oxford Univ. Press.
- Henrici, P. (1974). *Applied and Computational Complex Analysis*, 3 volumes. J. Wiley.
- Hoppensteadt, F.C. (2000). *Analysis and Simulations of Chaotic Systems*, Springer, New York.
- Hardy, G. (1910). *The orders of Infinity*. Cambridge Univ. Press.
- Kac, V., Cheung, P. (2002). *Quantum Calculus*. Springer, New York.
- Keller, J.B. (1985). *Lectures on Bifurcation Theory*. N. Y. Univ. 1965- (Notas não publicadas; reproduzidas nos textos da Escola de Mat. Aplicada, LNCC, Rio).
- Kline, M. (1974). *Mathematical Thought from Ancient to Modern Times*. Oxford U. Press.
- Knuth, D.E. (1998). Teach Calculus with big O, *Letter - Notices of the Am. Math. Soc.*, July, 687–688.
- Lax, P.D. (2000). *Functional Analysis*. J. Wiley.
- Leibniz, G.W. (1684). *Nova Methodus*. (trad. inglês em D.J. Struik-ed.-Source Book in Mathematics: 1200-1800. Harv. U. Press, 1969)
- Lutzen, J. (1982). *The Prehistory of the Theory of Distributions*. Springer-Verlag.
- Maslov, V.P., Fedoryuk, M.V. (1980). *Semi-Classical Methods*. Reidel.
- Mikusinski, J. (1958). *Operational Calculus*. Pergamon Press. (original publicado em 1953, Varsovia.)
- Nahin, P.J. (2002). *Oliver Heaviside*. J. Hopkins Univ. Press.
- Nazaikinskii, V.E., Shatalov, V.E., Sternin, B. Yu. (1998). *Methods of Noncommutative Analysis*. Walter de Gruyter, Berlin.
- Ruppert-Hall, A. (1980). *Philosophers at War*. Cambridge Univ. Press.
- Toeplitz, O. (2007). *The calculus: a genetic approach*. University of Chicago Press.
- Trefethen, L.N. (2005). *Spectra and Pseudospectra*. Princeton Univ. Press.
- Trenoguine, V.A. (1985). *Analyse Fonctionnelle*. MIR, Moscou.