

O conceito de medida, o continuum e o discreto

José Carlos Magossi 

Vania Rosa Figueiredo Izidoro 

Resumo

A palavra “medida” tem sido utilizada ao longo da história da humanidade em quase todos os setores da atividade humana. Não é de espantar que alterações no modo como se mede resultam em alterações no modo como a ciência e as tecnologias são desenvolvidas. Os refinamentos nos critérios de medida, num sentido amplo, ocorrem graças às tecnologias existentes, ou emergem de alguma regra bem definida, escrita em alguma linguagem, com algum fim científico ou prático. Enquanto, num passado remoto, media-se o diâmetro da Terra com base em semelhanças de triângulos, nos tempos atuais as tecnologias dão conta de tornar essas medidas muito mais precisas. Mesmo assim, um consenso ainda não é atingido, haja vista que o *continuum* matemático impõe restrições à realidade voltada às medições. Por exemplo, não há como em laboratórios utilizar π em sua plenitude, uma vez que aproximações são necessárias, levando-se em conta que é um número irracional com infinitas casas decimais. Essas incertezas, no quesito “precisão”, podem ser vistas como uma gangorra, em que de um lado há as medidas práticas da realidade em que vivemos e, de outro, as medidas teóricas. O objetivo neste artigo é, por um lado, expor, sob a ótica da matemática, alguns exemplos em que se caracterize a relação entre o *continuum* e o discreto, no quesito “medida”. Por outro, mostrar que essa relação pode indicar contradições, num palco de interações entre o mundo prático e o teórico, se não for feita uma leitura cuidadosa. Além de uma digressão histórica com exemplos, mostra-se que algo semelhante ocorre com o conceito de medida visto como quantidade de informação, denominado de entropia por C. E. Shannon. Há também cuidados a serem tomados com relação à entropia, vista sob a ótica de modelos discretos, e sua extensão para modelos contínuos, a entropia diferencial. Enquanto, do lado discreto, a quantidade de informação é positiva, a entropia diferencial, do lado contínuo, pode ser negativa, positiva ou arbitrariamente grande.

Palavras-chave: Medida, *continuum*, discreto, Shannon, entropia diferencial

Abstract

The word “measure” has been used throughout the history of humanity in almost every sector of human activity. It is not surprising that any changes in the way things are measured, also cause impacts in the development of science and technologies. The refinements in the measurement criteria, in a broad sense, occur thanks to existing technologies, or they emerge from some well-defined rule, written in some language, with some scientific or practical purpose. Whereas, in a remote past, the diameter of the Earth was measured based on similarities of triangles, in modern times technologies made these measures much more precise. Even so a consensus is not yet reached, given that the mathematical *continuum* imposes restrictions on measurement reality. For example,

there is no way to use π in its fullness in laboratories, since approximations are necessary, taking into account that it is an irrational number with infinite decimal places. This characterizes a seesaw, in which, on one hand, there are the practical measures in the reality we live in, and, on the other, the theoretical measures. The goal in this article is to expose that, on one hand, from the perspective of mathematics, some examples characterize the relation between *continuum* and the discrete, in the measured aspect. On the other hand, we show that this relationship can indicate contradictions, in a stage of interactions between the practical and the theoretical world, if no careful reading happens. Apart from a historical digression with examples, it is shown that something similar occurs with the concept of measure when it is seen as an amount of information, called entropy by C. E. Shannon. There is also care to be taken regarding entropy, seen from the point of view of discrete models, and their extension to continuous models, differential entropy. While on the discrete side the amount of information is positive, the differential entropy, on the continuous side, can be negative, positive or arbitrarily large.

Keywords: Measure, *continuum*, discrete, Shannon, differential entropy

1. Introdução

O conceito de medida relaciona-se à possibilidade de passar de um atributo que pode envolver continuidade e quantificadores a uma notação numérica que facilita, com base nas operações numéricas, transmitir informações acerca dos objetos investigados [34]. *Grosso modo*, medir é atribuir números (ou alguma grandeza, uma magnitude) a objetos seguindo determinadas regras. Em matemática o conceito de medida, extensivo a uma teoria da medida, de H. L. Lebesgue (1875-1941) [18], [5], relaciona-se à caracterização do conteúdo de um conjunto por um número real, que alicerça o desenvolvimento da teoria das probabilidades, por exemplo. Uma dificuldade inicial mostra-se quando se pretende elaborar uma definição de medida que seja extensiva o suficiente. Como construir relações entre objetos a serem medidos, com uma linguagem apropriada, de tal forma que haja uma exata interpretação numérica que possibilite leituras práticas de medidas em determinados contextos [30]. É bastante entusiástica a busca por um diálogo mais efetivo entre ciências empíricas e teorias de medida. Isso impacta a inserção de teorias de medidas em setores diversos, tais como educação, psicologia, ciências humanas, informática etc. [20, 27]. Neste artigo, o ponto de discussão caracteriza-se pela relação entre o conceito de medida quando aplicado a modelos discretos, e quando aplicado a modelos contínuos. Em muitos casos, utiliza-se o mesmo critério de medida, a mesma “régua”, seja para o caso discreto, seja para o caso contínuo. Por exemplo, ao somar $1 + \frac{1}{2} + \frac{1}{4} + \dots$, comumente se utiliza o nome “soma infinita”. No entanto, a palavra “soma” é reservada para conjuntos finitos de dados, sendo que a palavra “série” aplica-se aos casos em que uma quantidade infinita de parcelas é analisada. Ao escrever “soma infinita” emprega-se, de certa forma, a mesma “régua” para ambos os casos, o finito e o infinito. Caso uma distinção não seja feita, abre-se espaço para perguntas do tipo: “Como somar uma quantidade infinita de parcelas?”. De modo análogo, pode-se perguntar quais são as propriedades que se preservam (e quais não) ao se aplicar uma mesma propriedade de um conceito de medida em objetos oriundos de conjuntos discretos e também de conjuntos contínuos. A investigação dessas propriedades auxilia também na distinção entre as ferramentas que podem e as que não podem ser utilizadas com consistência para cada caso. As sutilezas que se escondem nas propriedades que residem, ou não, nos modelos contínuos e modelos discretos talvez possam ser mais bem compreendidas ao se examinar, com os devidos cuidados, o texto escrito no livro *Select Works of A. N. Kolmogorov*:

Se um mapa nos dá informações importantes sobre uma seção da superfície da

Terra, a microestrutura do papel e da tinta que ocorre no papel não tem nenhuma relação com a microestrutura da seção da superfície da Terra mostrada sobre esse papel. ([33], p.189) (Tradução dos autores).

Nesse caso, ao analisar o mapa de um país, exposto numa grande folha de papel, o uso de uma lente de aumento para descobrir, nas ranhuras do mapa, algumas propriedades das ruas de alguma cidade que consta no mapa não vai fornecer resultados consistentes. Nesse caso, a lente de aumento não é uma ferramenta apropriada para obter tais informações, haja vista que o mapa indica propriedades do país no sentido macroscópico e não microscópico. No entanto, algumas propriedades se preservam, seja no sentido macro como no micro, como, por exemplo, a fronteira entre as cidades.

Se o mundo macroscópico fornece (ou não) informações suficientes para uma análise precisa do mundo microscópico, esse é um assunto ainda deveras investigado [10, 11]. De certa forma, isso vem sendo investigado na matemática, tal como o problema da incomensurabilidade de $\sqrt{2}$, que tem sido discutido desde a Grécia antiga e, com base na sua incomensurabilidade, propiciado o surgimento de inúmeros paradoxos. Associar um número a um segmento de reta, tal como se faz nos dias de hoje, indicava, na Grécia antiga, na época de Euclides, ter de resolver, como caso particular, o problema da incomensurabilidade de $\sqrt{2}$. Para evitar esse problema, o geômetra Eudoxo opta por comparar figuras geométricas em vez de associar uma medida a elas. Por exemplo, se um quadrado tem lado a e outro quadrado tem lado β , então se diz que o quadrado de lado a é maior que o quadrado de lado β , em símbolos, $a^2 > \beta^2$, desde que o quadrado de lado β possa ser inserido no quadrado de lado a [19]. Essa ideia de medida procura, de certa forma, escapar às discussões sobre as quantidades infinitamente pequenas [2, 23]. Pode-se especular nesse caso que Eudoxo tenha evitado a relação entre os mundos microscópico e macroscópico e, tal como no exemplo acima, do mapa, tenha procurado por propriedades que fossem comuns a ambos os mundos.

Esse panorama só começa a mudar muito tempo depois, com o surgimento do conceito de limites por A. L. Cauchy (1789-1857) no século XIX [12],[13]. O conceito de limites indicou um divisor de águas para a matemática e acelerou inúmeros desenvolvimentos, tais como as alterações no conceito de medida. Assim se deu com a alteração do conceito de medida de áreas, volumes etc., voltado, na época anterior à de Cauchy, apenas às integrais de I. Newton (1643-1727) e G. Leibniz (1646-1716) [37, 22]. Analisar a classe de funções integráveis, que não sejam aquelas obtidas com base em antiderivadas (integrais de Newton-Leibniz) implica, de alguma forma, indicar um processo de medida voltado às partições do domínio da função investigada, no caso das integrais de Cauchy, Riemann, Kurzweil-Henstock, ou da partição do conjunto imagem da função, no caso das integrais de Lebesgue [6]. É certo que essas medições levam em conta o *continuum* matemático, mas é certo também que, com base nelas, desenvolvimentos matemáticos (e tecnológicos) puderam ser elaborados. Pode-se dizer que Cauchy desenvolveu ferramentas apropriadas para “medir” as quantidades infinitamente pequenas. Isto é, ele usou ferramentas adaptadas ao mundo microscópico. Casos clássicos da influência da teoria da medida e integração de Lebesgue, com ferramentas semelhantes, são: teoria das probabilidades, mecânica estatística, teoria da informação, termodinâmica, análise de Fourier etc. Longe de invocar o período de abstração matemática que caracterizou o início do século XX ([29], p.173), mas com o viés de que uma linguagem bem estruturada é condição *sine qua non* para sustentar inovações em matemática e tecnologia, entende-se, tal como ocorreu com a teoria da medida e integração de Lebesgue, que caracterizações do conceito de medida sejam desenvolvidas com fins de indicar diversificações na escolha de medidas e com isso possibilitar maior interação entre conceitos, evitando o medir com ferramentas inapropriadas. O importante com isso é diminuir as incertezas presentes na ciência com base no desenvolvimento de novos critérios, e novas tecnologias, associadas aos critérios de medida. As medidas fundadas na

geometria plana auxiliavam os trabalhos de Newton; com N. I. Lobachevsky (1792-1856) e C. F. Gauss (1777-1855), as da geometria não euclidiana auxiliaram os trabalhos de A. Einstein (1879-1955). As incertezas presentes na época de Newton não são as mesmas de hoje, haja vista que física clássica e física moderna são ramificações que indicam a evolução na diminuição dos níveis de incertezas. Conforme L. Brillouin (1889-1969) [10, 11], levando-se em conta a ideia de que um observador é que faz a medida, e em consonância com a ideia de entropia da termodinâmica, na física, a medida de distâncias muito pequenas é fisicamente impossível. Com isso, um ponto interessante é que, para o físico, no sentido operacional, as quantidades infinitamente pequenas (o *continuum*) são apenas abstrações, sem nenhum significado físico [10, 11]. Para Brillouin,

O matemático define o infinitamente pequeno, mas o físico é absolutamente incapaz de medi-lo; representa uma pura abstração sem nenhum significado físico. Se adotássemos o ponto de vista operacional, deveríamos eliminar o infinitamente pequeno das teorias físicas, mas, infelizmente, não temos ideia de como alcançar tal programa. ([10], p.x, Tradução dos autores.)

Não é objetivo trazer à tona as discussões de Brillouin sobre o determinismo ou sobre a impossibilidade de fazer medições para distâncias muito pequenas, mas sim indicar apenas um contexto histórico, uma trajetória de eventos e exemplos, com vistas a justificar a importância de leituras cuidadosas acerca do conceito de medida, nos modelos discretos e contínuos.

Mesmo com tantas incertezas e pluralidade no sentido das interações entre o que se mede e o que poderia ser medido, há problemas interessantes num domínio em que o conceito de medida relaciona-se com o conceito de incerteza, isto é, num domínio pertencente à teoria de Shannon, ou teoria da informação [14]. C. E. Shannon (1916-2001), em seu artigo de 1948 [31], para seu modelo matemático de comunicação, indicou uma medida para a quantidade de informação a ser transmitida entre um emissor X e um receptor Y , através de um canal de comunicação. Com essa medida, denominada por ele de entropia¹, é possível tratar informações e melhorar a eficiência na transmissão de dados em sistemas de comunicação [26, 4, 14]. Se $X = \{x_1, x_2, \dots, x_i, \dots, x_m\}$ e $Y = \{y_1, y_2, \dots, y_j, \dots, y_n\}$ são variáveis aleatórias discretas, e $p(x_i)$ é a probabilidade *a priori* de que $X = x_i$, então a entropia discreta de X , para um sistema completo de probabilidades,

$$H(X) = - \sum_{i=1}^M p(x_i) \log p(x_i),$$

mede a quantidade de informação da experiência representada pela variável X . A quantidade de informação *a posteriori* é a incerteza eliminada *a priori*. No caso de variáveis aleatórias contínuas, se $f(x)$ representa uma função de densidade de probabilidade da variável aleatória absolutamente contínua X , então a entropia diferencial de X ,

$$h(X) = - \int_{-\infty}^{+\infty} f(x) \log f(x) dx,$$

representa, não um número, uma quantidade de informação associada à variável aleatória X , como na entropia para o caso discreto, mas uma medida relativa de incerteza [24] ou uma mudança na incerteza². Com base na medida da quantidade de informação, entropia no sentido de Shannon,

¹Shannon adotou o mesmo nome utilizado na termodinâmica por R. Clausius (1822-1888) [21].

²Note-se, nesse caso, que essa medida fundamenta-se no conceito de integral, haja vista que a existência de $h(X)$ está condicionada à existência da integral presente em sua definição [14].

uma pergunta interessante, em consonância com a discussão aqui apresentada sobre as leituras distintas entre o mundo microscópico e mundo macroscópico, é: Como se dá a relação entre entropia discreta e entropia diferencial? [31, 4, 26].

Esse conceito, de entropia diferencial, que pode ser visto como uma nova leitura da medida da informação para o caso contínuo, tem aplicações em campos além da mecânica estatística e teoria da comunicação, tais como finanças, econometria, processamento de imagens, ecologia, bioestatística etc. [24]. Ou seja, há desafios matemáticos interessantes relacionados à medida de uma quantidade de informação, relacionados ao passar de modelos discretos para modelos contínuos [26]. Mais interessante é que a “régua” utilizada para medir informação para modelos discretos (entropia discreta) é diferente daquela utilizada para medir informações em modelos contínuos.

Na seqüência, neste texto, expõem-se, no sentido histórico, exemplos de medidas em modelos contínuos que, quando lidos sob a ótica de modelos discretos, podem incorrer em propensos paradoxos ou contradições. No sentido físico, ilustra-se, de modo breve, a rica discussão, presente em [10, 11], acerca da impossibilidade de medir distâncias muito pequenas. Nesse caso, abre-se espaço para a percepção da importância das ferramentas matemáticas no desenvolvimento de modelos e sua imediata limitação nos modelos do mundo em que habitamos. Expõe-se também o avanço científico que se obtém quando alterações nos modelos contínuos produzem novos desenvolvimentos na matemática, tal como se deu com o conceito de integrais. Finaliza-se o artigo com o conceito de medida relacionado às comunicações. Nesse caso, há uma medida da quantidade de informação (ou incerteza *a priori*), no caso de modelos discretos, mas há um campo novo de investigação que reside na quantidade de informação, incerteza, presente nos modelos contínuos (entropia diferencial). Isso não deixa de ser algo conhecido na história da humanidade, reforçado neste artigo com alguns exemplos, sobre qual é o divisor de águas, se é que existe, entre as ferramentas matemáticas adaptadas para modelos discretos e aquelas adaptadas para os contínuos.

2. Possíveis equívocos com a palavra “medida”

Nesta seção alguns exemplos são expostos com o objetivo de mostrar o quanto de cuidado deve-se ter ao utilizar ferramentas desenvolvidas para modelos discretos com fins de obter resultados em modelos contínuos, ou lançar mão de ferramentas desenvolvidas para modelos contínuos na busca por resultados em modelos discretos. Não que isso não possa ser feito, mas precaução talvez seja a palavra correta, pois contradições podem emergir da leitura, talvez descuidada, da relação entre contínuo e discreto, como a não enumerabilidade de \mathbb{R} , incomensurabilidade e conjuntos infinitos. Note-se com os exemplos seguintes que o conceito de medida deve ser pautado em firmes definições, para que contradições não emergjam de leituras não atentas.

2.1. Área infinita e volume finito

Considere-se o cálculo da área da superfície de revolução obtida ao girar a função $y = \frac{1}{x}$ limitada no intervalo $[1, +\infty)$, em torno do eixo x . A fórmula que indica o cálculo da área da superfície desse sólido, obtida por revolução ([17], p.410) em torno do eixo x , é:

$$\text{Área} = \int_1^{+\infty} 2\pi f(x) \sqrt{1 + (f'(x))^2} dx.$$

Se $f(x) = \frac{1}{x}$, então $f'(x) = \frac{-1}{x^2}$, e daí:

$$\begin{aligned} \text{Área} &= \int_1^{+\infty} 2\pi f(x) \sqrt{1 + \left(\frac{-1}{x^2}\right)^2} dx = \int_1^{+\infty} 2\pi \frac{1}{x} \frac{\sqrt{x^4 + 1}}{x^2} dx. \\ &= \lim_{R \rightarrow +\infty} \int_1^R 2\pi \frac{\sqrt{x^4 + 1}}{x^3} dx. \end{aligned}$$

Essa integral imprópria não converge; logo, infere-se que a área da superfície de revolução em questão é infinita. Com o auxílio de computadores é possível verificar, para $R = 10^n$ e $n \in \mathbb{N}$, que,

$$\varphi(n) = \int_1^{10^n} 2\pi \frac{\sqrt{x^4 + 1}}{x^3} dx \tag{1}$$

$$= \left(\sqrt{2} - \ln(1 + \sqrt{2}) + \operatorname{arcsenh}(100^n) - 100^{-n} \sqrt{10000^n + 1} \right) \pi. \tag{2}$$

Então $\varphi(1) = 15,177$, $\varphi(2) = 29,645$, $\varphi(4) = 58,58$, $\varphi(8) = 116,45$, $\varphi(12) = 174,32$, $\varphi(15) = 217,72$ etc. Ou seja, quanto maior é o valor de n , maior é o valor da integral e, por conseguinte, da área associada à integral. Nesse caso, tem-se que

$$\lim_{n \rightarrow +\infty} \varphi(n) = +\infty.$$

Por outro lado, o cálculo do volume desse mesmo sólido, obtido por revolução da função $\frac{1}{x}$ em torno do eixo x , no intervalo $[1, +\infty)$, é feito ([17], p.381) pela seguinte integral:

$$\text{Volume} = \int_1^{+\infty} \pi f(x)^2 dx = \int_1^{+\infty} \pi \left(\frac{1}{x}\right)^2 dx = \int_1^{+\infty} \left(\frac{\pi}{x^2}\right) dx.$$

Essa é uma integral imprópria, que converge, ou seja,

$$\text{Volume} = \lim_{R \rightarrow +\infty} \int_1^R \pi x^{-2} dx = \lim_{R \rightarrow +\infty} \left[\frac{-\pi}{x} \right]_1^R = \lim_{R \rightarrow +\infty} \left[\frac{-\pi}{R} + \frac{\pi}{1} \right] = \pi.$$

Nesse caso, a unidade do valor do volume igual a π vai depender da unidade que possa ser utilizada na descrição do intervalo $[1, +\infty)$. O volume poderá ser expresso em m^3 , cm^3 etc. Note-se, então, que a mesma figura, o sólido de revolução com área infinita se $R \rightarrow +\infty$, conforme pode ser observado na figura 1, tem volume finito, se $R \rightarrow +\infty$. Essa propensa “contradição” advém de leituras discretas em problemas cuja essência volta-se a modelos contínuos. Cabe observar que, em tese, não há contradição, mas a ideia de área infinita e volume finito, para um mesmo sólido³, pode gerar confusão haja vista não ser algo tão intuitivo. Uma leitura descuidada pode caracterizar que isso seja uma contradição.

³Nesse caso considera-se que o sólido existe no intervalo $[1, +\infty)$, algo distante da realidade prática em que vivemos.

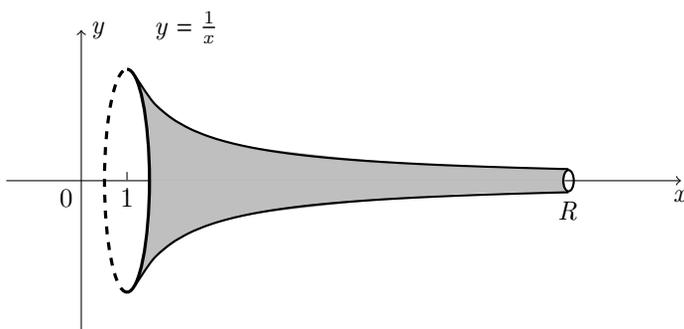


Figura 1: Revolução de $y = \frac{1}{x}$ em torno do eixo x .

2.2. A roda de Aristóteles

Supõe-se a existência de duas circunferências, C_1 e C_2 , de tal forma que a circunferência C_2 seja maior que a circunferência C_1 . Supõe-se ainda que ambas sejam colocadas uma dentro da outra, isto é, a circunferência C_1 está dentro da circunferência C_2 , de tal forma que possam ambas deslizar sobre retas u e v , paralelas entre si. A circunferência C_1 desliza sobre a reta u e a C_2 , sobre a reta v . Note-se que cada ponto de uma circunferência está em correspondência biunívoca com cada ponto da outra, conforme pode ser observado na figura 2. No entanto, ambas percorrem o mesmo trajeto de A até B , no mesmo tempo. Se r_2 é o raio da circunferência C_2 e r_1 é o raio da circunferência C_1 , então o comprimento da circunferência C_1 é $2\pi r_1$, e o comprimento da circunferência C_2 é $2\pi r_2$.

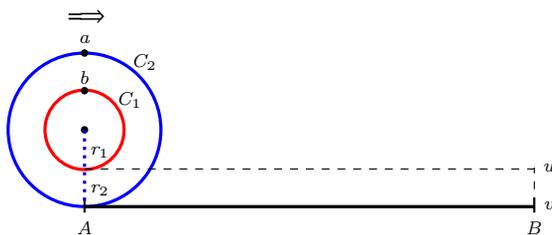


Figura 2: Roda de Aristóteles

Mas, por hipótese, $r_1 < r_2$, então $2\pi r_1 < 2\pi r_2$, isto é, o comprimento da circunferência C_1 é menor do que o comprimento de C_2 . Mas, levando-se em conta o desenho, tem-se a impressão de que a trajetória de A até B refere-se ao comprimento de cada circunferência. Se assim for, tem-se então uma possível contradição. No entanto a curva cicloide realizada pela circunferência menor (com base no ponto b e reta suporte u) não é a mesma curva cicloide que a realizada pela circunferência maior (com base no ponto a e reta suporte u), conforme pode ser observado na curva cicloide da figura 3. Note-se que o comprimento do segmento contido na reta suporte u , nesse caso, é $2\pi r_1$.

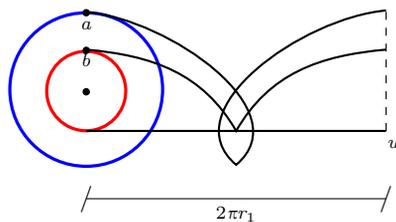


Figura 3: Curva cicloide com início em b sobre a reta u

De modo análogo, a curva cicloide realizada pela circunferência maior (com base no ponto a e reta suporte v) não é a mesma curva cicloide que a realizada pela circunferência menor (com base no ponto b e reta suporte v), conforme pode ser observado na figura 4. Note-se que o comprimento do segmento contido na reta suporte v , nesse caso, é $2\pi r_2$.

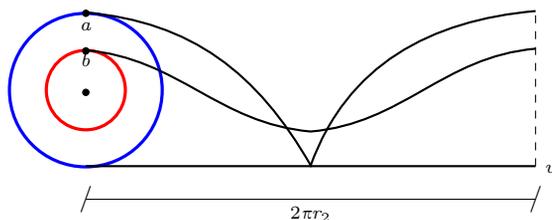


Figura 4: Curva cicloide com início em a sobre a reta v

Assim, conforme for a rotação da roda, o tamanho da reta suporte dependerá do ponto de partida do trajeto. Caso o trajeto seja determinado pela circunferência menor a partir do ponto b , ao longo da reta u , então o percurso é menor que o trajeto iniciado no ponto a , com base na circunferência maior ao longo da reta v . No artigo *Aristotle's Wheel: Notes on the History of a Paradox* [16] encontra-se uma rica discussão sobre esse tema. Além disso, o fato de ambos os conjuntos terem uma correspondência biunívoca (entre cada ponto de C_1 e C_2) não garante que tenham trajetórias idênticas.

2.3. Aquiles e a tartaruga

Na literatura, atribui-se a Zenão de Eleia, da Grécia antiga, um paradoxo acerca de uma corrida entre o guerreiro Aquiles e uma tartaruga. Nessa corrida, a tartaruga tem uma vantagem sobre Aquiles, pois está à sua frente antes do início da corrida. A ideia de Zenão é, com essa simulação, mostrar que Aquiles nunca venceria a corrida, pois nunca alcançaria a tartaruga. Quando Aquiles chegar ao ponto em que a tartaruga saiu (dada a vantagem inicial), ela já terá percorrido um certo trajeto, e estará num ponto, digamos P_1 . Quando Aquiles chegar a P_1 , a tartaruga já estará em outro ponto, P_2 , e assim sucessivamente. Se isso se repetir *ad infinitum*, então Aquiles nunca alcançará a tartaruga. Esse é o paradoxo de Zenão de Eleia, século V antes de Cristo, na Grécia antiga. De certa forma, pode-se inferir que Zenão discutia acerca das quantidades infinitamente pequenas [23] e das contradições que emergem de sua utilização, como no caso de Aquiles e a tartaruga. Se um objeto desloca-se de um ponto A até um ponto B , esse objeto precisa passar

pela metade do caminho, diga-se C , que une A a B . De modo análogo, precisa passar pela metade do caminho que une C até o ponto final B . E assim sucessivamente. Como, argumenta Zenão, há infinitos pontos a percorrer, o objeto nunca chegará ao ponto B . Ou seja, se a distância de A até B é de u metros, então para ir de A até B é preciso percorrer as distâncias, medidas em metros,

$$\frac{u}{2} + \frac{u}{4} + \frac{u}{8} + \dots$$

Essa soma é uma série geométrica de razão $\frac{1}{2}$ e converge para u . Essa convergência é uma conclusão com base na matemática e na teoria de limites, em que quantidades infinitamente pequenas são levadas em conta. Não é uma conclusão no sentido intuitivo, mas no sentido formal fundamentado numa teoria de limites e convergência, desenvolvido muitos anos após a conhecida estória de Zenão. No entanto Zenão, sem o conhecimento de limites (desenvolvido por Cauchy em 1821), levantava discussões acerca dessas quantidades e de sua conexão com o mundo em que vivemos. Nesse caso, pode-se citar A. Alexander, que, de certa forma, traz à luz do artigo um raciocínio elaborado por Zenão:

Mas se os paradoxos de Zenão e o problema da incomensurabilidade provam alguma coisa, é que o sonho de um encaixe perfeito entre matemática e o mundo físico é insustentável. Na escala dos infinitamente pequenos, números não correspondem a objetos físicos, e qualquer tentativa de forçar esse encaixe conduz a paradoxos e contradições. ([2], pp.19-20)

Nesse caso, nota-se que a palavra “soma” destina-se a operações com um número finito de parcelas. A palavra “série” refere-se ao processo em que há um número infinito de parcelas. Assim, “soma infinita” é uma expressão que pode gerar dúvidas, pois devemos lembrar a utilização de operações finitárias, em cenários não finitários.

2.4. Medidas de ângulos: graus e radianos

As medidas de ângulos em graus e em radianos são muito utilizadas. A medição em radianos é uma das medidas que muitas vezes geram confusão nos meios escolares. Um grau é $\frac{1}{360}$ de uma circunferência, isto é, uma parte de 360 numa circunferência. Um radiano é o ângulo central de uma circunferência unitária C , conforme pode ser visto na figura 5, que corresponde a um arco, cujo comprimento é igual ao raio unitário da circunferência C . A relação entre ângulos em graus e radianos é dada pela seguinte expressão:

$$2\pi \text{ radianos} \approx 360^\circ.$$

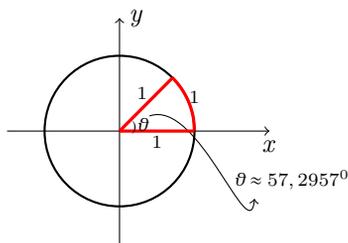


Figura 5: Ângulo θ equivalente a 1 radiano

Note-se nessa expressão que, de um lado, há o número irracional π , e do outro, o número inteiro 360° . Isso não deixa de ser uma relação entre contínuo e discreto. Nesse caso,

$$1 \text{ radiano} \approx 57,2957^\circ.$$

No ensino fundamental e médio, nos tópicos sobre trigonometria e medida de ângulos, é comum utilizar medidas em radianos em vez de medidas em graus⁴. Conforme pode ser observado no artigo [35], as medidas de ângulos em graus e radianos impactam o conceito de derivadas. Ou seja, se a medida de ângulos for em graus, então a derivada da função $\text{sen } x$ será $0,01753 \cos x$. Por outro lado, se a medida de ângulos for em radianos, então a derivada da função $\text{sen } x$ será a função $\cos x$. Isso ocorre, pois, para que a derivada da função seno seja a função cosseno, o limite de $\frac{\text{sen } x}{x}$ deve ser 1 quando x tende a zero. Mas, no caso de medida em graus, tem-se que:

$$\lim_{x \rightarrow 0} \frac{\text{sen } x}{x} = \frac{\pi}{180} = 0,017453 \dots$$

3. Medidas (e não medidas) para modelos contínuos

Nesta seção, expõem-se três exemplos. Os dois primeiros exemplos de medidas para modelos contínuos são o problema de Plateau e a medida do calor em uma barra metálica. O de Plateau é um problema com simples enunciado, de fácil intuição, no entanto de difícil solução, tanto é que sua solução está relacionada à primeira *medalha Fields* em matemática⁵. O problema de condução de calor numa barra metálica é também de “natureza” contínua e exibido no presente artigo, pois sua solução, exposta no início do século XIX, foi considerada formalmente correta, mas inadequada, haja vista não haver na época de sua solução ferramentas matemáticas sólidas o suficiente para garantir que as hipóteses da solução, proposta por J. Fourier (1768-1830) em 1807, fossem confirmadas. Nesses dois exemplos, ferramentas de modelos contínuos foram utilizadas para sua solução, quais sejam, equações diferenciais. O terceiro exemplo mostra a impossibilidade, no sentido físico, da implementação do conceito de medida em sua forma geral, que, de certa maneira, coloca um limite no “medir”. Nesse caso, mostra-se a dificuldade em relacionar o contínuo com o discreto.

⁴Em alguns casos, utiliza-se a medida de ângulos em grados, que é a centésima parte de um ângulo reto. Isto é, uma circunferência dividida em 400 partes.

⁵J. Douglas (1897-1865) e L. V. Ahlfors (1907-1996) foram os primeiros ganhadores da medalha Fields em 1936. <https://www.mathunion.org/imu-awards/fields-medal>. Acesso em 8 de junho de 2021, às 17h25.

3.1. O problema de Plateau

O problema de Plateau [3] é o de encontrar uma superfície, delimitada por uma curva num espaço de três dimensões, com a menor medida de área possível. Esse problema deve-se ao físico belga J. A. F. Plateau (1801-1883), que mostrou, em 1849, que uma superfície mínima pode ser obtida quando se mergulha, por exemplo, um arame numa vasilha com água e sabão. A superfície S que se forma com a água e o sabão é a superfície com a menor área possível tendo o arame como fronteira, conforme pode ser observado na figura 6. Essa típica experiência com água e sabão, que pode ser estendida para outras formas de fronteira (que não seja apenas um arame como uma curva simples fechada), produz uma superfície em três dimensões que é a superfície de área mínima que pode ser obtida tendo o arame como fronteira.

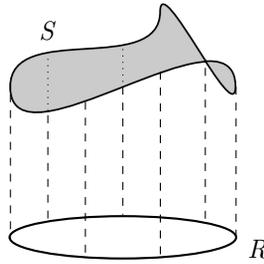


Figura 6: Exemplo de uma superfície de Plateau

Esse tipo de problema data da época de L. Euler (1707-1783) e J. L. Lagrange (1736-1813), e hoje se insere na área do Cálculo de Variações. Esse é um problema de medida que tem uma explicação intuitiva e pode ser observado de modo muito simples via experiências práticas. No entanto a demonstração da veracidade dessas experiências práticas não é trivial, muito menos simples. Ela foi exposta pelo matemático americano J. Douglas [15], e de modo independente pelo matemático húngaro-americano T. Radó (1895-1965) [25]. No caso do problema de Plateau, a medida mínima é obtida levando-se em conta o *continuum* real; caso contrário, num modelo discreto, a estimativa pode se resumir apenas à visualização da referida superfície e à intuição de que seja mínima.

3.2. Medida da temperatura em uma barra metálica

Com fins de resolver o problema de condução de calor em uma barra metálica⁶, Fourier indica, em 1807, uma solução matemática formal para a equação diferencial parcial $a^2 u_{xx} = u_t$, cuja resposta [9] é:

$$u(x, t) = \sum_{n=1}^{+\infty} b_n e^{-\frac{n^2 \pi^2 a^2 t}{L^2}} \sin\left(\frac{n \pi x}{L}\right).$$

O objetivo com sua solução é medir a temperatura $u(x, t)$ da barra em qualquer instante $t > 0$ do tempo na posição x . Considera-se que a barra tem comprimento $0 \leq x \leq L$, e leva-se em conta que $u(0, t) = 0$ e $u(L, t) = 0$ (a temperatura nas extremidades da barra mantém-se a zero grau). Considera-se ainda que existe uma distribuição inicial de temperatura na barra, $f(x) = u(x, 0)$ para

⁶Problema proposto pelo Institut de France, Paris, que conferiria um prêmio a quem fornecesse a melhor solução.

$0 \leq x \leq L$, conforme pode ser observado na figura 7. Para Fourier, em sua solução, as condições iniciais deveriam satisfazer:

$$f(x) = \sum_{n=1}^{+\infty} b_n \sin\left(\frac{n\pi x}{L}\right) \quad \text{para} \quad b_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx.$$

Sua solução não convenceu o júri encarregado de analisar os trabalhos que concorriam ao referido prêmio, pois, segundo eles, havia dúvidas acerca das funções que poderiam ser escritas como séries infinitas de senos e cossenos. Essas dúvidas seriam resolvidas a partir do momento em que se entendessem com mais precisão as ideias de convergência em séries e integrais. Com os desenvolvimentos de Lebesgue, em 1902, sobre um novo conceito de integral fundamentado numa teoria da medida, a solução formal de Fourier pôde ser analisada [36]. Com ela, os problemas de Fourier (fundamentados em convergências de séries e integrais) foram resolvidos. Note-se que um problema prático de condução do calor que pode ser verificado ao esquentar uma barra metálica, tal como o cabo de uma panela que vai ao fogo, precisou de quase cem anos para que sua solução fosse comprovada. Isso se deve ao fato de Lebesgue ter introduzido medidas para o trato com o *continuum* real.

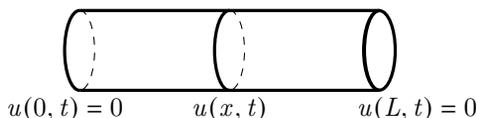


Figura 7: Barra metálica

3.3. Incerteza no medir com “régua”

Quantidades infinitamente pequenas fazem parte do dia a dia do trabalho do matemático. Aproximações são elaboradas e processos de convergência são investigados. O mesmo não se dá no mundo físico em que vivemos, haja vista que aproximações são utilizadas e análises dos erros são levadas em conta. A migração de modelos discretos para modelos contínuos, e vice-versa, implica análises cuidadosas acerca das ferramentas utilizadas e das aproximações, principalmente se o foco reside no conceito de medida.

No que concerne a discussões acerca da impossibilidade de medir distâncias muito pequenas, Brillouin ([11], p.32) indica a impossibilidade de, por exemplo, medir uma distância muito menor do que 10^{-15} cm, pois não há nenhuma “régua” disponível para medir distâncias tão pequenas. Suponha-se, segundo Brillouin, que se queira medir algo da ordem de 10^{-50} cm. O único comprimento padrão que possibilitaria levar adiante tal medida, como comparação de magnitudes, conforme consta em [11], seria o comprimento λ de uma onda eletromagnética. Note-se que, quanto menor for o comprimento de onda (maior frequência), maior será a precisão, no entanto maior será a energia cedida pela radiação em virtude da relação de M. Planck (1858-1947) entre energia e frequência da radiação

$$E = h.v,$$

em que E é a energia do fóton, h é a constante de Planck ($6,62607015 \times 10^{-34} J.s$) e v é a frequência da radiação. Assim, para tal medida, levando-se em conta a comparação com o comprimento de

onda λ , tem-se:

$$E = hv = h \cdot \frac{c}{\lambda} \approx \frac{2 \cdot 10^{-16}}{10^{-50}} \approx 2 \cdot 10^{34} \text{ C.G.S.}$$

Ainda segundo Brillouin [11], à página 32, isso representa uma quantidade de energia suficiente para destruir todo o laboratório em que a medida seria realizada, e também o planeta Terra. Desse modo, é possível inferir que, se algo não é possível de ser medido, então não deve ser levado em conta nas experiências laboratoriais. Seguindo essa linha de pensamento, pode-se conjecturar que no trabalho do matemático há estruturas em que quantidades infinitamente pequenas são utilizadas, quantidades essas que não deixam de ser abstrações que correspondem a situações físicas impossíveis ([10], p.236). Por outro lado, não há como não dizer que essas quantidades infinitamente pequenas sejam parte de desenvolvimentos matemáticos que acabam por auxiliar na criação de tecnologia,

O matemático define cuidadosamente os números irracionais. O físico nunca encontra esses tais números. Não importa se o que ele mede é representado por um número finito, com muitas figuras, e com uma certa quantidade de incerteza. O matemático estremece com a incerteza e tenta ignorar os erros experimentais. ([11], p.33) (Tradução dos autores).

Ainda segundo Brillouin, para o físico essas quantidades são apenas abstrações, as quais estão muito além das realizações laboratoriais. Num laboratório, a menor distância que pode ser observada é aquela que pode ser medida de acordo com a quantidade de energia disponível nesse laboratório ([10], p.236).

Desse modo, uma linguagem bem definida, bem estruturada, seja no sentido abstrato, seja no laboratorial, é parte integrante da ciência e, por mais que haja incerteza, tanto na linguagem abstrata como na laboratorial, elas acabam por se complementar. Pode-se inferir que o avanço da ciência implica a diminuição das incertezas. No caso do presente artigo, a diminuição das incertezas referentes ao conceito de medida.

4. Integrais: medir intervalos de números reais

O conceito de medida está presente na matemática dos intervalos de números reais. Com vistas a expor a relevância do conceito de medida nos intervalos de números reais, expõem-se algumas mudanças significativas que ocorreram quando pequenas alterações no trato com os intervalos de números reais indicaram alterações no desenvolvimento do conceito de integrais⁷. Nesse caso, alterações foram desenvolvidas no processo de medição com ferramentas de modelos contínuos (sentido microscópico). Há outros exemplos da matemática que podem igualmente ser mencionados, tais como os voltados à topologia, aos intervalos encaixantes, aos intervalos de Cantor etc. Nem sempre há uma designação clara acerca de como medir, nem há uma unidade representativa, haja vista que normalmente a magnitude associada à medida é um número puro, sem unidades.

4.1. Cauchy

Cauchy, em sua definição de integrais, utilizou uma partição de um intervalo $[a, b]$. Se $I = [a, b]$ é um intervalo fechado limitado em \mathbb{R} , então uma **partição** de I é um conjunto finito, ordenado,

⁷Nessa seção segue-se o exposto no artigo [22], bem como os livros [6, 7].

$\mathcal{P} = \{x_0, x_1, \dots, x_{n-1}, x_n\}$ de pontos em I , tal que $a = x_0 < x_1 < x_2 < \dots < x_n = b$ ([7], pp.199-200). Os pontos de \mathcal{P} são usados para dividir $I = [a, b]$ em subintervalos fechados que não se sobrepõem,

$$I_1 = [x_0, x_1], I_2 = [x_1, x_2], \dots, I_n = [x_{n-1}, x_n],$$

cuja união é $[a, b]$. Representa-se uma partição de $[a, b]$ por

$$\mathcal{P} = \{[x_{i-1}, x_i]\}_{i=1}^n.$$

Cauchy elaborou somas, para cada partição, de modo a obter a área da região limitada entre função $f(x)$ e o eixo x . Se $\mathcal{P} = \{[x_{i-1}, x_i]\}_{i=1}^n$ é uma partição de $[a, b]$, então a **soma de Cauchy** de uma função contínua $f : [a, b] \rightarrow \mathbb{R}$ correspondente a \mathcal{P} é o número

$$S(f; \mathcal{P}) = \sum_{i=1}^n f(x_{i-1})(x_i - x_{i-1}).$$

Para cada retângulo que será parte da área da região sob a curva $f(x)$, Cauchy utilizou $(x_i - x_{i-1})$ como base do retângulo e lançou mão de $f(x_{i-1})$ (os pontos x_{i-1} à esquerda de cada subintervalo I_i da partição de $[a, b]$) como sua altura. Pode-se escrever que Cauchy utilizou uma “régua”, conforme pode ser observado na figura 8, com uma escala limitada. Com base nisso, escreve que uma função contínua $f : [a, b] \rightarrow \mathbb{R}$ é dita ser **Cauchy integrável** em $[a, b]$, se existe um número $L \in \mathbb{R}$, em símbolos $C \int_a^b f(x) dx = L$, tal que para cada $\varepsilon > 0$ existe um $\delta > 0$ tal que se \mathcal{P} é uma partição de $[a, b]$ com $|\mathcal{P}| < \delta$, então

$$|S(f; \mathcal{P}) - L| < \varepsilon.$$

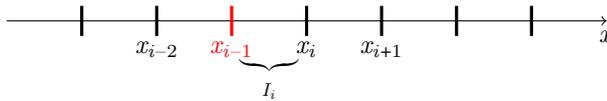


Figura 8: “Régua” de Cauchy

4.2. Riemann

As somas de Cauchy consideram, em cada subintervalo de uma partição, o valor da esquerda, x_{i-1} , do subintervalo $[x_{i-1}, x_i]$ como argumento da função f . Por outro lado, B. Riemann (1826-1866) aprofunda os estudos sobre a classe de funções integráveis e, em vez de escolher, em cada partição, o valor x_{i-1} (tal como fez Cauchy), escolhe um valor qualquer, ou seja, considera um valor $t_i \in [x_{i-1}, x_i]$. A “régua” de Riemann, conforme pode ser observado na figura 9, é, com abuso de linguagem, bem semelhante às que utilizamos no dia a dia.

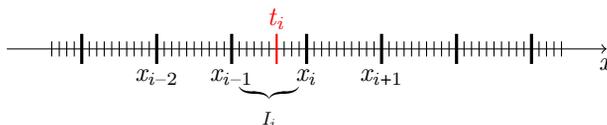


Figura 9: “Régua” de Riemann

Essa flexibilidade permite estabelecer condições necessárias e suficientes para a existência de uma integral. Riemann aprimorou a régua utilizada por Cauchy e em sua régua escalas mais diversificadas eram permitidas, quaisquer pontos num intervalo podem ser utilizados. Se $\mathcal{P} = \{[x_{i-1}, x_i]\}_{i=1}^n$ é uma partição do intervalo $I = [a, b]$, uma **etiqueta** é um ponto $t_i \in I_i = [x_{i-1}, x_i]$, para $i = 1, 2, \dots, n$. Uma **partição etiquetada** de $I = [a, b]$ é um conjunto de pares ordenados

$$\dot{\mathcal{P}} = \{([x_{i-1}, x_i], t_i)\}_{i=1}^n$$

de subintervalos I_i e suas respectivas etiquetas t_i . Se $\dot{\mathcal{P}} = \{([x_{i-1}, x_i], t_i)\}_{i=1}^n$ é uma partição etiquetada, então a **soma de Riemann** de uma função $f : [a, b] \rightarrow \mathbb{R}$ correspondente a $\dot{\mathcal{P}}$ é o número

$$S(f; \dot{\mathcal{P}}) = \sum_{i=1}^n f(t_i)(x_i - x_{i-1}).$$

Uma função $f : [a, b] \rightarrow \mathbb{R}$ é dita ser **Riemann integrável** em $[a, b]$ se existe um número $L \in \mathbb{R}$, em símbolos $\mathcal{R} \int_a^b f(x) dx = L$, tal que, para cada $\varepsilon > 0$, existe um $\delta > 0$ tal que, se $\dot{\mathcal{P}}$ é uma partição etiquetada de $[a, b]$ com $\|\dot{\mathcal{P}}\| < \delta_\varepsilon$, então

$$|S(f; \dot{\mathcal{P}}) - L| < \varepsilon.$$

4.3. Kurzweil e Henstock

Ainda com fins de aumentar a classe de funções integráveis, e quiçá inovar nas possibilidades de medir todos os subconjuntos de \mathbb{R} [32], J. Kurzweil e R. Henstock (1923-2007) desenvolvem uma “régua”, ao estilo de Riemann, mas com escalas flexíveis. Nessa nova “régua”, que inclui a “régua” de Riemann, as escalas de medida podem ser ajustadas de acordo com o que se vai medir. Para isso utiliza o termo calibre [7]. Uma função **calibre** em um intervalo $I = [a, b]$ é uma função estritamente positiva definida em I . Se δ é um calibre em I , então uma partição **etiquetada** $\dot{\mathcal{P}}$ de $[a, b]$, em que $I_i = [x_{i-1}, x_i]$ e t_i é uma etiqueta de I_i , é dita ser δ -**fina** se

$$t_i \in I_i \subseteq [t_i - \delta(t_i), t_i + \delta(t_i)] \text{ para } i = 1, 2, \dots, n.$$

Uma função $f : [a, b] \rightarrow \mathbb{R}$ é dita ser **Kurzweil-Henstock integrável** [6, 7], em $[a, b]$ se existe um número $L \in \mathbb{R}$, em símbolos $\mathcal{KH} \int_a^b f(x) dx = L$, tal que, para cada $\varepsilon > 0$, existe um calibre $\delta_\varepsilon > 0$ em $[a, b]$, tal que, se $\dot{\mathcal{P}}$ é qualquer partição δ_ε -fina de $[a, b]$, então

$$|S(f; \dot{\mathcal{P}}) - L| < \varepsilon.$$

Do mesmo modo que, com simples alterações na “régua” utilizada por Cauchy, Riemann expandiu sua classe de funções integráveis, as integrais de Kurzweil-Henstock foram expandidas graças também a simples alterações na “régua” utilizada por Riemann, tal como pode ser observado na figura 10. Nela nota-se que, de acordo com o calibre δ_ε , as escalas na régua alteram-se em tamanho.

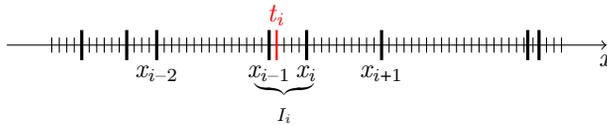


Figura 10: “Régua” de Kurzweil-Henstock

Note-se a importância que o conceito de medida tem no desenvolvimento da matemática, haja vista que pequenas alterações na “régua” utilizada para “medir” intervalos de números reais acarretaram profundas alterações na classe das funções integráveis. Há ainda as medidas (e teoria da medida) de Lebesgue [18, 5], não tratadas neste artigo, que de modo análogo impactaram significativamente muitos ramos da matemática.

O objetivo da discussão sobre as medidas em intervalos de números reais e o conceito de integrais relacionam-se também com a próxima seção, que indicará como o conceito de medida pode também ser caracterizado como medida de quantidade de informação. Como será exibido na próxima seção, a diferença entre medir informação para modelos discretos e modelos contínuos reside na utilização, *grosso modo*, do sinal de somatório \sum para modelos discretos e do sinal de integral \int para modelos contínuos. Ou seja, a existência de medidas de informação para o caso contínuo envolve a existência de integrais, e, por conseguinte, a existência de funções integráveis. Ou seja, a “ponte” entre o discreto e o contínuo, no quesito “medida de informação”, passa pelo conceito de integral.

5. Medidas de informação discreta e contínua

É bem plausível escrever que a palavra “medir” esteja relacionada à utilização de alguma régua, de alguma escala numérica, para medir alguma coisa. É o caso também da medida de informação, ou medidas de incerteza, desenvolvidas por Shannon [31]. Shannon indicou uma medida de informação, denominou-a de entropia, para um modelo matemático de comunicação, a qual é essencial na caracterização da era digital, nos tempos atuais. Os processos de comunicação valem-se de sinais, contínuos ou discretos, que são transmitidos (codificados e decodificados) por canais de comunicação (fios, antenas, cabos etc). Shannon tratou da comunicação entre dois pontos levando em conta a existência de ruídos (interferências no canal de comunicação). Os modelos de comunicação, no sentido da entropia de Shannon, podem ser lidos de acordo com os conceitos da área de engenharia, levando-se em conta sinais, recepção, emissão, redundância, codificação, decodificação, filtro, canais etc.

Ao se levar em conta que a entropia “mede” a quantidade de informação, então essa propensa “régua” precisa de ajustes ao lidar com medidas em modelos discretos e contínuos. Para modelos discretos tem-se a entropia discreta, para modelos contínuos tem-se a entropia diferencial. Note-se que, nos modelos contínuos, a entropia diferencial de Shannon não compartilha das mesmas leituras que a dos modelos discretos. Ou seja, há propriedades da entropia discreta que não valem na entropia diferencial. *Grosso modo*, no caso da medida de informação, entropia discreta e diferencial não compartilham da mesma “régua”. No entanto, ao se fazer uma extensão para o caso de duas variáveis aleatórias contínuas, X e Y , então a medida da informação mútua $I(X; Y)$ compartilha de propriedades comuns ao caso discreto ([26], p.269), ([4], p.236) e [24]. Ou seja, há propriedades da entropia diferencial que valem para o caso discreto, mas não valem para o caso contínuo, e há propriedades da informação mútua que valem para ambas. Isso acaba por

possibilitar outras leituras, e aplicações, da entropia diferencial, em áreas que não a da teoria da informação [24].

5.1. Modelos discretos

Em seu artigo de 1948 [31], Shannon utilizou probabilidades para caracterizar uma medida de incerteza ou, como é comumente escrito, uma medida de informação. Shannon trabalhou com a estrutura da informação como sendo derivada de considerações estatísticas, não levando em conta aspectos semânticos, valores ou significados relacionados à palavra informação. Quanto maior a incerteza *a priori* associada a um experimento, maior será a quantidade de informação (incerteza eliminada após o experimento). A seguinte fórmula, para símbolos $x_1, x_2, \dots, x_i, \dots, x_m$ com respectivas probabilidades *a priori*

$$p(x_1) = p_1, p(x_2) = p_2, \dots, p(x_i) = p_i, \dots, p(x_m) = p_m,$$

representa a quantidade de informação por símbolo:

$$H(p_1, p_2, \dots, p_m) = -K \sum_{i=1}^m p_i \ln p_i.$$

A constante K determina a unidade dessa informação. Se o alfabeto é binário, então $K = \frac{1}{\ln 2} = \log_2 e$, e daí se pode escrever que a quantidade de informação é medida em bits.

$$\begin{aligned} H(p_1, p_2, \dots, p_m) &= -\frac{1}{\ln 2} \sum_{i=1}^m p_i \ln p_i = -\log_2 e \sum_{i=1}^m p_i \left(\frac{\log_2 p_i}{\log_2 e} \right) \\ &= -\sum_{i=1}^m p_i \log_2 p_i \quad \frac{\text{bits}}{\text{símbolo}}. \end{aligned}$$

Se cada símbolo consome um tempo t em segundos, então se tem a medida em $\frac{\text{bits}}{\text{segundos}}$.

No caso em que se tem eventos equiprováveis $p_1 = p_2 = \frac{1}{2}$, tal como no lançamento de uma moeda “honestas”, então

$$H(p_1, p_2) = -\sum_{i=1}^2 p_i \log_2 p_i = -\sum_{i=1}^2 \frac{1}{2} \log_2 \frac{1}{2} = -\left(\frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{2} \log_2 \frac{1}{2} \right) = 1.$$

Ou seja, a medida da quantidade de informação no lançamento de uma moeda é 1 bit.

Shannon generalizou sua fórmula de entropia para variáveis aleatórias discretas,

$$X = \{x_1, x_2, \dots, x_i, \dots, x_m\},$$

com $p(x_i) = p_i$ para $1 \leq i \leq m$ e $\sum_{i=1}^m p_i = 1$. Para um alfabeto binário,

$$H(X) = -\sum_{i=1}^m p(x_i) \log_2 p(x_i)$$

é a fórmula de **entropia discreta** de Shannon, e denota uma medida de quantidade de informação.

5.2. Modelos contínuos

A transmissão de informações pode se dar por meio de símbolos discretos, ou por sinais contínuos, ou por ondas elétricas, ou como uma função contínua em relação ao tempo, estabelecida num intervalo $t \in [a, b]$. Nesse intervalo a amplitude do sinal assume valores contínuos com base numa função de densidade de probabilidade ([26], p.267). Tal como em outros setores da matemática, a mudança de discreto para contínuo exige cuidado. No caso da fórmula de entropia para modelos discretos, $H(X) = -\sum_{i=1}^m p(x_i) \log_2 p(x_i)$, sua extensão para o caso contínuo é dada pela **entropia diferencial**, em que $f(x)$ é uma função densidade de probabilidade,

$$h(X) = - \int_{-\infty}^{+\infty} f(x) \log_2 f(x) dx.$$

Em consonância com o que tem sido exposto neste artigo, $H(X)$ tem propriedades que $h(X)$ não tem. Ou seja, enquanto $H(X)$ é positiva, $h(X)$ pode ser negativa (“informação negativa”), infinita (“quantidade infinita de informação”) ou mesmo positiva. Isso indica que, para modelos contínuos, $h(X)$ não pode ser considerada, de imediato, como uma medida de informação, a menos que restrições apropriadas sejam indicadas [28].

5.3. Entropia diferencial $h(X)$ negativa

Nesta seção o objetivo é exibir um exemplo em que a entropia diferencial pode ser negativa. No caso em que X é uma variável aleatória discreta que pode assumir os valores x_1, x_2, \dots, x_n , com respectivas probabilidades $p(x_1) = p_1, p(x_2) = p_2, \dots, p(x_n) = p_n$, tal que $\sum_{i=1}^n p_i = 1$, a entropia $H(X) = -\sum_{i=1}^n p_i \log_2 p_i$ é sempre um número não negativo, haja vista que a função $x \log_2 x$ é negativa para valores entre 0 e 1, conforme pode ser observado na figura 11.

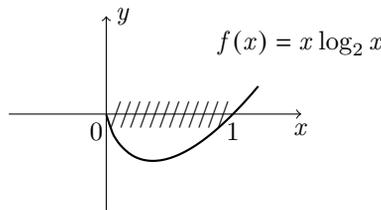


Figura 11: Função logarítmica $y = x \log_2 x$

Já para o caso de uma variável aleatória contínua X , com função densidade de probabilidade $f(x)$, a seguinte condição deve ser satisfeita:

$$\int_{-\infty}^{+\infty} f(x) dx = 1.$$

Essa condição é análoga à condição $\sum_i p_i = 1$ em variáveis aleatórias discretas. A construção de exemplos em que se tenha a entropia negativa leva em conta a determinação de uma função

$f(x) \geq 0$, cuja área no plano xy seja 1, de acordo com a variação de alguns parâmetros. Por exemplo, para números reais $a, h \geq 0$, seja $f(x) = \frac{h}{a}x$ se $0 \leq x \leq a$, conforme pode ser observado na figura 12. Note-se que os infinitos valores de a e h que satisfazem $\frac{ah}{2} = 1$ implicam que a área hachurada A será igual a 1 (como é o caso quando $a = 1$ e $h = 2$). Entende-se então que há valores de a e h tais que $\frac{ah}{2} = 1$ e que, ao assumir, como hipótese, que a área da região A vale 1, tem-se então que $A = \frac{ah}{2} = 1$.

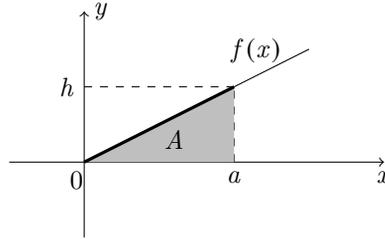


Figura 12: Função densidade de probabilidade $f(x) = \frac{hx}{a}$.

Com base nessa definição da função f , mostra-se que ela é uma função densidade de probabilidade, ao mostrar que $\int_{-\infty}^{+\infty} f(x) = 1$ e que $f(x) \geq 0$. A função f é maior ou igual a zero por construção. Resta então mostrar que

$$\int_{-\infty}^{+\infty} f(x) = 1.$$

$$\int_{-\infty}^{+\infty} f(x) dx = \int_0^a \frac{hx}{a} dx = \left[\frac{hx^2}{2a} \right]_0^a = \left[\frac{ha^2}{2a} - 0 \right] = \frac{ha}{2}.$$

Como, por hipótese, a área $A = \frac{ah}{2} = 1$, tem-se então que

$$\int_{-\infty}^{+\infty} f(x) dx = \frac{ah}{2} = 1,$$

e $f(x)$ é uma função densidade de probabilidade.

Com base nessa função $f(x)$, que é uma função densidade de probabilidade, torna-se possível determinar o valor de $h(X)$. Assim, ao considerar, sem perdas de generalidade, que o logaritmo é expresso na base e , tem-se:

$$h(X) = - \int_{-\infty}^{+\infty} f(x) \log f(x) dx = - \int_{-\infty}^{+\infty} \frac{h}{a} x \log \frac{h}{a} x dx.$$

Ao fazer a substituição $u = \frac{hx}{a}$, tem-se que $du = \frac{h}{a} dx$. Logo,

$$\begin{aligned}
 h(X) &= - \int_{-\infty}^{+\infty} u \log u \left(\frac{h}{a} \right) du = - \frac{h}{a} \int_{-\infty}^{+\infty} u \log u du = - \frac{h}{a} \int_0^a u \log u du. \\
 &= \frac{h}{a} \left[\frac{u^2}{2} \cdot \log u - \frac{u^2}{4} \right]_0^a = \frac{h}{a} \left[\frac{a^2}{2} \cdot \log a - \frac{a^2}{4} - 0 \right] = \frac{ah}{2} \log a - \frac{ah}{4}.
 \end{aligned}$$

Portanto,

$$h(X) = \frac{ah}{2} \left[\log a - \frac{1}{2} \right].$$

Como, por hipótese, tem-se que $\frac{ah}{2} = 1$, então o valor de $\log a - \frac{1}{2}$ determinará o valor de $h(X)$. Por exemplo, se $a = \sqrt{2}$, então

$$h(X) = \frac{ah}{2} [\log a - 1] = 1 \cdot [\log \sqrt{2} - 1] = 0,3465 - 0,5 = -0,1534 < 0.$$

Ou seja, $h(X) < 0$.

Assim, levando-se em conta que o logaritmo é expresso na base e , tem-se:

- Se $a = \sqrt{e}$, então $h(X) = 0$.

$$\log a - \frac{1}{2} = 0 \Rightarrow \log a = \frac{1}{2} \Rightarrow a = e^{\frac{1}{2}} \Rightarrow a = \sqrt{e}.$$

- Se $a < \sqrt{e}$, então $h(X) > 0$.

$$\log a - \frac{1}{2} < 0 \Rightarrow \log a < \frac{1}{2} \Rightarrow a < e^{\frac{1}{2}} \Rightarrow a < \sqrt{e}.$$

- Se $a > \sqrt{e}$, então $h(X) < 0$.

$$\log a - \frac{1}{2} > 0 \Rightarrow \log a > \frac{1}{2} \Rightarrow a > e^{\frac{1}{2}} \Rightarrow a > \sqrt{e}.$$

A entropia diferencial $h(X)$ pode ser nula, positiva ou negativa. Ou seja, enquanto no caso discreto a quantidade de informação $H(X)$ é um número positivo ou nulo, no caso contínuo $h(X)$ pode ser nula, positiva ou negativa. Isso mostra que a “medida de informação” para o caso contínuo é diferente da medida de informação para o caso discreto.

5.4. A informação mútua $I(X; Y)$.

Nesta seção, sem pretensões de completude na exposição, expõe-se, com base em referências bibliográficas, o conceito de informação mútua. Para o caso discreto, tal como pode ser observado em [26], [4] ou [14], a informação mútua pode ser expressa como

$$I(X; Y) = H(X) - H(X|Y).$$

Nesse caso, tanto $H(X)$ quanto a entropia condicional $H(X|Y)$ são positivas ou nulas. Daí, $I(X; Y) \geq 0$ para o caso discreto. Para o caso contínuo, a informação mútua $I(X; Y)$ tem propriedades semelhantes à da entropia discreta, isto é, é maior ou igual a zero. Para verificar essa condição, consideram-se duas variáveis aleatórias contínuas, a variável aleatória contínua X e a variável aleatória contínua Y . Considera-se também uma função densidade de probabilidade conjunta $f(x, y)$, tal que $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$. Assim, conforme pode ser observado em [26], à

página 276, tem-se que:

$$\begin{aligned}
 I(X; Y) &= - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \log \frac{f(x, y)}{f_X(x)f_Y(y)} dx dy \\
 I(X; Y) &= - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \log \frac{f_X(x)}{f(x|y)} dx dy \\
 &\geq - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \left[\frac{f_X(x)}{f(x|y)} - 1 \right] \log e dx dy \\
 &= - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_Y(y)f_X(x) \log e dx dy + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \log e dx dy \\
 &= 1.1. \log e - \log e = 0.
 \end{aligned}$$

Portanto, $I(X; Y) \geq 0$.

Desse modo, mostrou-se que, enquanto a entropia diferencial $h(X)$ e a entropia discreta $H(X)$ sofrem alterações na medida da quantidade de informação quando se alteram os modelos de discreto para contínuo, a informação mútua, isto é, a informação que transita pelo canal de comunicação, se mantém-se inalterada. Ou seja, a medida de ambas é não negativa. Essa é uma propriedade que reside na intersecção dos modelos discretos e contínuos.

6. Conclusões

Na matemática, vista como uma linguagem composta de estruturas e objetos [17], não somente abstrações são possíveis, mas também uma abertura total para processos imaginativos que permitem a criação de estruturas adaptadas aos, por exemplo, processos de medição. Isso não assegura que todos esses processos tenham uma contraparte presente na realidade do mundo em que vivemos. Ou seja, não há uma regra específica que indique quais devam ser os desenvolvimentos que não sejam apenas uma multiplicação de estruturas matemáticas, e quais os que impactam a construção de novas tecnologias. Não há também como indicar impedimentos nesses desenvolvimentos. O conceito de medida, a noção intuitiva de medir alguma coisa, apresenta-se na história da humanidade e vem se desenvolvendo ao longo dos tempos. Medidas na época de G. Galilei (1564-1642) são muito distintas das medidas nos tempos atuais. Isso não significa que se está no ápice do estado da arte dos processos de medição. Tal como indicado por Brillouin, a menor distância possível a ser medida depende da quantidade de energia à disposição no laboratório em que a medida será realizada. Isso revela que novos desenvolvimentos podem ser feitos de modo a diminuir as incertezas, aumentar as precisões no quesito “medida”, como por exemplo, a clássica e fantástica história da definição da medida “metro”, na França, em 1792 [1], que impactou os sistemas de medida ao redor do mundo. Neste artigo, exemplos foram exibidos com vistas a mostrar que há oscilação, em relação ao conceito de medida, entre as leituras advindas de ferramentas discretas para modelos contínuos e leituras advindas de ferramentas contínuas para modelos discretos. Mostrou-se que há casos em que, mesmo que a leitura seja descuidada, o resultado é equivalente, haja vista a utilização de uma ferramenta que serve aos dois propósitos, quais sejam, a medida em modelos discretos e nos contínuos.

Há também contradições, ao se limitar a leitura ao “mundo discreto” quando o objetivo é observar o “mundo contínuo”. Como exemplo, pode-se citar a importância de ter leituras atentas a propriedades advindas do *continuum* real. Atenção deve-se ter, por exemplo, para não concluir que

o conjunto \mathbb{R} é “maior” que um intervalo dele mesmo⁸. Isso pode levar a contradições, se uma atenção maior não for empregada à distinção entre o todo e a parte. Ou seja, esse é um exemplo de uma leitura do mundo macroscópico acerca de modelos do mundo microscópico. Há também, não abordada neste artigo, a teoria da medida de Lebesgue, presente no início do século XX, e também todas as discussões acerca de medidas (não de Lebesgue), voltadas às ciências humanas, que elencam critérios de medida apropriados a cada cenário em particular. Note-se, com base na figura 13, a existência de alguns pontos-chave que impactaram (e podem ainda impactar) o conceito de medida ao longo da história da matemática.

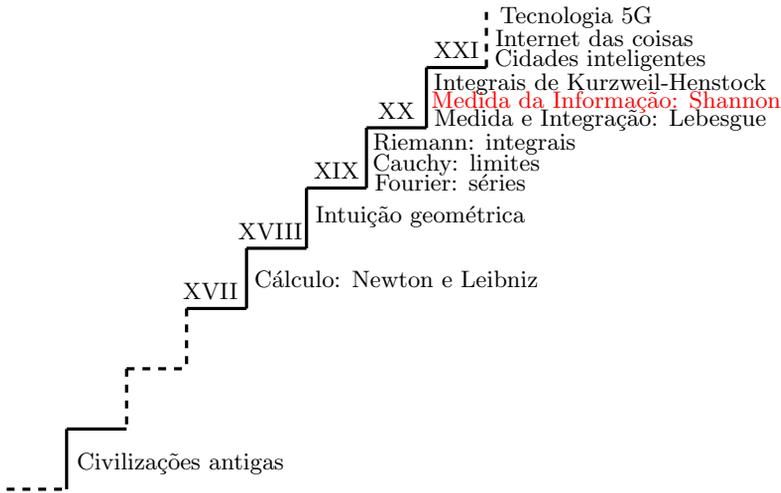


Figura 13: Evolução do conceito de medida ao longo dos séculos.

Estima-se, com este artigo, que a proposta da análise do conceito de medida em patamares discretos e contínuos, e da análise de quais são as propriedades que pertencem a ambos os patamares, possa incentivar novos desenvolvimentos e auxiliar na elucidação de equívocos, ou aparentes contradições. Em particular, estima-se que novos desenvolvimentos possam surgir das investigações acerca das medidas de informação, e das leituras operacionais possíveis, no sentido das propriedades pertencentes aos modelos discretos e contínuos. Note-se que desenvolvimentos nessa área impactam os desenvolvimentos tecnológicos que podem propiciar melhorias nas comunicações e no desenvolvimento humano, tais como tecnologia 5G, internet das coisas, cidades inteligentes etc.

Referências

- [1] Alder, K. *Mesurer le monde: L’incroyable histoire de l’invention du mètre*. Traduction: Martine Devillers-Argouac’h. Éditeur Flammarion, 2015.
- [2] Alexander, A. *Infinitesimal - A teoria matemática que mudou o mundo*. Jorge Zahar Editor Ltda, Rio de Janeiro, 2014.
- [3] Almgren, F. J. *Plateau’s Problem: An Invitation to Varifold Geometry*. American Mathematical Society (First Edition, 1966). UK Edition, 2001.

⁸Nesse caso, a função $f : \mathbb{R} \rightarrow (-1, 1)$, tal que $f(x) = \frac{x}{1+|x|}$, é uma função bijetora, ou seja, tanto \mathbb{R} quanto $(-1, 1)$ têm o mesmo número de elementos, conforme exposto nos trabalhos de G. Cantor sobre teoria de conjuntos [8].

- [4] Ash, R. B. *Information Theory*. Dover Publications, New York, 1990. (Originalmente publicado pela Interscience Publishers, New York, 1965).
- [5] Bartle, R. G. *The Elements of Integration and Lebesgue Measure*. John Wiley & Sons, Inc., New York, 1966.
- [6] Bartle, R. G. “A Modern Theory of Integration”. *Graduate Studies in Mathematics*, Volume 32. American Mathematical Society, Providence, Rhode Island, 2001.
- [7] Bartle, R.G.; Sherbert, D.R. *Introduction to Real Analysis*, Fourth Edition, John Wiley & Sons, New York, 2011.
- [8] Belna, J. P. *Cantor*. Figuras do Saber. Tradução Guilherme João de Freitas Teixeira, Revisão Técnica Michel Paty. Editora Estação Liberdade, São Paulo, 2011.
- [9] Boyce, W. E.; DiPrima, R. C. *Equações diferenciais elementares e problemas de valores de contorno*. LTC- Editora, sexta edição. Rio de Janeiro, 1998.
- [10] Brillouin, L. *Science and information theory*. Dover edition, first published in 1935 (is an unabridged republication of the second edition (1962) of the work originally published in 1956 by Academic Press, Inc, New York).
- [11] Brillouin, L. *Scientific Uncertainty and information*. Academic Press Inc. New York, 1964.
- [12] Cauchy, A. L. *Cours d'analyse (Cours d'analyse de l'école royale polytechnique. 1^{re} partie: analyse algébrique*. Paris, 1821.
- [13] Cauchy, A. L. *Résumé des leçons données à l'école royale polytechnique sur le calcul infinitésimal*. Vol. 1. Imprimerie royale, 1823.
- [14] Cover, T. M.; Joy A. T. *Elements of Information Theory*, John Wiley & Sons, 1991.
- [15] Douglas, J. *Solution of the problem of Plateau*. Transactions of the American Mathematical Society. Published by American Mathematical Society , Jan., 1931, Vol. 33, n° 1, pp. 263-321, 1931.
- [16] Drabkin, I. E. *Aristotle's Wheel: Notes on the History of a Paradox*. Osiris, Vol. 9, pp. 162-198. The University of Chicago Press on behalf of The History of Science Society. 1950.
- [17] Lang, S. *A First Course in Calculus*. Fifth Edition. Springer-Verlag, New York, 1986.
- [18] Lebesgue, H. L. *Intégrale, longueur, aire*. Annali di Mathematica Pura ed Applicata (3), 7: 231-359, 1902.
- [19] Lintz, R. G. *História da Matemática*. v. I. Blumenau: EdFURB, 1999.
- [20] Luce, R. D. *The Ongoing Dialog between Empirical Science and Measurement Theory*. Journal of mathematical psychology 40, 7898, 1996.
- [21] Magossi, J. C.; Paviotti, J. R. “Incerteza em Entropia”. *Revista Brasileira de História da Ciência*. Rio de Janeiro, V.12, n° 1, pp.84-96, janeiro-junho, 2019.
- [22] Magossi, J. C.; Barros, A. C. C. “O conceito de Integrais nos cursos de Cálculo”. *Revista do Professor de Matemática Online*. Sociedade Brasileira de Matemática. PMO, pp.239-262, v.7, n° 2, 2019.
- [23] Magossi, J. C. “O sonho de Lagrange”. *Professor de Matemática Online – PMO*, Sociedade Brasileira de Matemática, v.8, n° 1, pp.43-63, 2020.
- [24] Michalowicz, J. V.; Nichols, J. M.; Bucholtz, F. *Handbook of differential entropy*. CRC Press, Taylor & Francis Group, A Chapman & Hall Book. New York, 2014.

- [25] Rado, T. *On Plateau's Problem*. Published by Mathematics Department, Princeton University. Annals of Mathematics, Jul., 1930, Second Series, Vol. 31, No. 3, pp.457-469, 1930.
- [26] Reza, F. M. *An Introduction to Information Theory*. McGraw-Hill, New York, 1961.
- [27] Rossi, G. B. *Measurement and Probability: a probabilistic theory of measurement with applications*. Springer Series in Measurement Science and Technology. Springer, Dordrecht, 2014.
- [28] Rioul, O. *Teoria da Informação e da Codificação*. Editora da Unicamp e Editora Universidade de Brasília. Campinas, 2018. (Tradução do original em francês por José Carlos Magossi).
- [29] Salsburg, D. *Uma senhora toma chá... como a estatística revolucionou a ciência no século XX*. Jorge Zahar Editor Ltda, Rio de Janeiro, RJ, 2009.
- [30] Scott, D.; Suppes, P. "Foundational Aspects of Theories of Measurement. Association for Symbolic Logic". *The Journal of Symbolic Logic*, Vol. 23, N° 2, pp.113-128, Jun., 1958.
- [31] Shannon, C. E. "A mathematical theory of communication". *Bell System Technical Journal*, v. 27, pp.379-423, 623-656, 1948.
- [32] Shenitzerand, A.; Steprian, S. *The evolution of integration*. The American Mathematical Monthly, Vol. 101, No. 1 (Jan., 1994), pp.66-72
- [33] Shiriyayev, A. N. *Select works of A. N. Kolmogorov. Volume III. Information theory and the Theory of Algorithms*. Springer-Science+Business Media, B.V. Dordrecht, 1993.
- [34] Volken, H. "Les fondements : mesure pour mesure". *Revue européenne des sciences sociales*. Tome XLV, n° 138, pp.41-53, 2007. (*European Journal of Social sciences*). <http://journals.openedition.org/ress/192;DOI:10.4000/ress.192>
- [35] Watanabe, R. G. "Seno de 30 é um meio ?" *Revista do Professor de Matemática – RPM*. Número 30, pp.26-32, São Paulo, 1996.
- [36] Wilcox, H. J.; Myers, D. L. *An Introduction to Lebesgue Integration and Fourier Series*. Dover Publications, New York, 1978.
- [37] Yee, L. P.; Vyborny, R. *The Integral: An Easy Approach after Kurzweil and Henstock*. Cambridge University Press, Cambridge, 2000.

José Carlos Magossi
Universidade Estadual de Campinas
<magossi@unicamp.br>

Vania Rosa Figueiredo Izidoro
Centro Estadual de Educação Tecnológica Paula Souza - Nova Odessa
<vania.izidoro@etec.sp.gov.br>

Recebido: 10/09/2021
Publicado: 17/08/2022