International Journal of Mathematical Education in Science and Technology, Vol. 44, No. 4, 2013, 595–609
(author's reprint)

Zero-bounded limits as a special case of the squeeze theorem for evaluating single-variable and multivariable limits

Eleftherios Gkioulekas*

University of Texas-Pan American, Edinburg, TX 78539-2999
(*Received 04 Jan 2012, Accepted 24 Sep 2012, Published online: 19 Nov 2012*)

Many limits, typically taught as examples of applying the "squeeze" theorem, can be evaluated more easily using the proposed zero-bounded limit theorem. The theorem applies to functions defined as a product of a factor going to zero and a factor that remains bounded in some neighborhood of the limit. This technique is immensely useful both for single-variable limits and multidimensional limits. A comprehensive treatment of multidimensional limits and continuity is also outlined.

1. Introduction

In most Calculus courses, it is customary to have a section where the concept of the squeeze theorem is introduced and then used to evaluate limits. Because a proper solution requires understanding neighborhoods and the establishment of two inequalities, giving a lower and upper bound, homework problems in this area can be challenging for students. Consequently, instructors sometimes decide that it is best to underemphasize or skip the topic altogether. As a matter of fact, most squeeze theorem limit problems fall under the special category of a zero-bounded limit. A zero-bounded limit is one in which the function can be broken into a product of two functions where one function converges to zero and the other function is bounded. If we show that a limit is zero-bounded, then the zero-bounded limit theorem implies that the limit goes to zero.

The main advantage of this zero-bounded limit theorem is that it requires only the establishment of one inequality, and because the properties of absolute values can be always leveraged to prove that one inequality, the argument tends to be easier. Other advantages of teaching students the zero-bounded limit theorem are that: (1) it is easy to identify zero-bounded limits as such merely by inspection; (2) it is a great opportunity to reinforce the teaching of the manipulation of absolute value inequalities, a skill very important in more advanced courses like analysis; (3) the same technique can immensely simplify the calculation of two-dimensional or three-dimensional limits often encountered in multivariable calculus. Finally, it is important to underscore that many of these limits cannot be evaluated via other techniques such as the L'Hospital theorem.

In this paper, we present the zero-bounded limit theorem and show how to apply it to both single-variable and multi-variable limits. The presentation of the theory and solved examples will also illustrate the relevance and convenience of quantifier notation to representing definitions and mathematical arguments presented in calculus and analysis coursework. The paper is organized as follows. In section 2 we derive the zero-bounded theorem for the single-variable case, and show how to use

*Corresponding author. Email: drlf@hushmail.com

ISSN: 1045-1129 print/ISSN 1029-0346 online

DOI: 10.1080/0020739X.2012.742148
http://www.informaworld.com

it on specific examples. In section 3 we extend the zero-bounded limit theorem to multi-variable limits. Finally, in section 4 we illustrate how continuity and the zero-bounded limit theorem can be combined to handle more challenging multivariable limits.

2. Single-variable zero-bounded limits

Let $f : A \to \mathbb{R}$ be a function with $A \subseteq \mathbb{R}$ and consider the limit $x \to \sigma$ with σ an accumulation point of the domain A. In general, possible assignments for σ are $\sigma = x_0$, or $\sigma = x_0^+$,or $\sigma = x_0^-$ with $x_0 \in \mathbb{R}$, or $\sigma = +\infty$, or $\sigma = -\infty$. Given a $\delta > 0$, for each of these possibilities we define the corresponding *generalized neighborhood* $N(\sigma, \delta)$ as:

$$N(\sigma,\delta) = \begin{cases} (x_0-\delta, x_0+\delta) - \{x_0\} & , \text{ if } \sigma = x_0 \\ (x_0-\delta, x_0) & , \text{ if } \sigma = x_0^- \\ (x_0, x_0+\delta) & , \text{ if } \sigma = x_0^+ \\ (1/\delta, +\infty) & , \text{ if } \sigma = +\infty \\ (-\infty, -1/\delta) & , \text{ if } \sigma = -\infty \end{cases} . \tag{1}$$

We will use this notation throughout this paper. For the cases $\sigma = x_0$, $\sigma = x_0^+$, and $\sigma = x_0^-$, $N(\sigma, \delta)$ coincides with the definition of the *deleted neighborhood*, but we have generalized the definition to also account for the cases $\sigma = \pm\infty$. Note that for all cases, a smaller δ implies a tighter neighborhood.

We also use the following definition for *accumulation points* and the limit of a function converging to a number:

$$\sigma \text{ accumulation point of } A \iff \forall\ \delta > 0 : N(\sigma,\delta) \cap A \neq \emptyset,$$

$$\lim_{x\to\sigma} f(x) = \ell \iff \begin{cases} \sigma \text{ accumulation point of } A \\ \forall\ \varepsilon > 0 : \exists\ \delta > 0 : \forall\ x \in A \cap N(\sigma,\delta) : |f(x) - \ell| < \varepsilon. \end{cases}$$

Note that the brace in the expression above will represent the conjunction (i.e. the logical "and") between two or more statements, with each statement given on a separate line. Furthermore, the generalized neighborhoods satisfy the following obvious properties:

$$0 < \delta_1 \leq \delta_2 \Longrightarrow N(\sigma,\delta_1) \subseteq N(\sigma,\delta_2), \tag{2}$$

$$N(\sigma,\delta_1) \cap N(\sigma,\delta_2) = N(\sigma, \min\{\delta_1,\delta_2\}), \tag{3}$$

which can be used to derive the general properties of limits from the definition given above.

Although the zero-bounded theorem can be derived from the squeeze theorem, the most efficient way to develop the theory is by deriving both the squeeze theorem and the zero-bounded limit theorem from a preliminary lemma, which is, in and of itself, useful for evaluating simple limits. Note that most textbooks [1–5] prefer to give a direct proof of the squeeze theorem. An indirect proof similar to the one given here was also given by Apostol [6], although he did not explicitly display the intermediate lemma. The lemma and its proof are as follows:

Lemma 2.1: *Let $f : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$ be two functions, and let σ be an* *accumulation point of A. Then,*

$$\left.\begin{array}{l}\exists\ \delta > 0 : \forall\ x \in N(\sigma,\delta) : |f(x)| \leq g(x)\\ \lim\limits_{x\to\sigma} g(x) = 0\end{array}\right\} \Longrightarrow \lim_{x\to\sigma} f(x) = 0.$$

Proof: Let $\varepsilon > 0$ be given. From the hypothesis:

$$\lim_{x\to\sigma} g(x) = 0 \Longrightarrow \exists\ \delta > 0 :\ \forall\ x \in A \cap N(\sigma,\delta) : |g(x)| < \varepsilon.$$

Let $x \in A \cap N(\sigma,\delta)$ be given. Then:

$$\begin{aligned} g(x) \geq |f(x)| \geq 0 &\Longrightarrow g(x) = |g(x)| \\ \Longrightarrow |f(x)| \leq g(x) = |g(x)| < \varepsilon &\Longrightarrow\ |f(x)| < \varepsilon. \end{aligned}$$

It follows that

$$\begin{aligned} &[\forall\ \varepsilon > 0 :\ \exists\ \delta > 0 :\ \forall\ x \in A \cap N(\sigma,\delta) : |f(x)| < \varepsilon] \\ \Longrightarrow &\lim_{x\to\sigma} f(x) = 0. \end{aligned}$$

□

The squeeze theorem is derived from this lemma in Appendix A, since it is not the main concern of this paper. To formulate the zero-bounded limit theorem, we begin by stating the definition of a bounded function:

Definition 2.2: Let $f : A \to \mathbb{R}$ be a function, and let $B \subseteq A$. Then,

$$f \text{ is bounded on } B \Longleftrightarrow\ \exists\ a > 0 :\ \forall\ x \in B : |f(x)| \leq a.$$

Then, the zero-bounded limit theorem states that:

Theorem 2.3: *Let $g : A \to \mathbb{R}$ and $b : A \to \mathbb{R}$ be two functions, and let σ be an accumulation point of A. Then,*

$$\left.\begin{array}{l}\exists\ \delta > 0 : b \textit{ is bounded on } N(\sigma,\delta)\\ \lim\limits_{x\to\sigma} g(x) = 0\end{array}\right\} \Longrightarrow \lim_{x\to\sigma}[b(x)g(x)] = 0.$$

Proof: Since b is bounded on $N(\sigma,\delta)$, we have

$$\exists\ a > 0 : \forall\ x \in N(\sigma,\delta) : |b(x)| \leq a.$$

It follows that

$$\forall\ x \in N(\sigma,\delta) : |b(x)g(x)| = |b(x)||g(x)| \leq a|g(x)|.$$

We also note that

$$\lim_{x\to\sigma} g(x) = 0 \Longrightarrow \lim_{x\to\sigma} |g(x)| = 0 \Longrightarrow \lim_{x\to\sigma} a|g(x)| = 0.$$

From the above two equations, via the Lemma 2.1, we have $\lim_{x\to\sigma} b(x)g(x) = 0$.
$\square$

To apply the theorem, it is sufficient to establish that $b(x)$ is bounded on a generalized neighborhood $N(\sigma, \delta)$ of the limit $x \to \sigma$ in which δ can be chosen to be as small as is necessary. In most zero-bounded limits involving trigonometric functions we use the bounds

$$\begin{aligned} &\forall\, x \in \mathbb{R} : |\sin x| \leq 1, \\ &\forall\, x \in \mathbb{R} : |\cos x| \leq 1, \end{aligned}$$

in conjunction with the following properties of the absolute values

$$\begin{aligned} &\forall\, x, y \in \mathbb{R} : |x + y| \leq |x| + |y|, \\ &\forall\, x, y \in \mathbb{R} : |x - y| \leq |x| + |y|, \\ &\forall\, x, y \in \mathbb{R} : |xy| = |x||y|, \\ &\forall\, x \in \mathbb{R} : \forall\, y \in \mathbb{R} - \{0\} : |x/y| = |x|/|y|. \end{aligned}$$

We illustrate the technique with the following two examples:

Example 2.4 Evaluate the following limit

$$\lim_{x\to+\infty} \frac{\sin x \cos x - 3\sin(2x)}{x^2+1}.$$

First, we can see by inspection that the given function can be broken into a zero-bounded product as follows:

$$f(x) = \frac{\sin x \cos x - 3\sin(2x)}{x^2+1} = [\sin x \cos x - 3\sin(2x)]\frac{1}{x^2+1} = b(x)g(x),$$

with $b(x) = \sin x \cos x - 3\sin(2x)$ and $g(x) = 1/(x^2+1)$. We may thus proceed with the following solution.

Solution : Define $b(x) = \sin x \cos x - 3\sin(2x)$. Since

$$\begin{aligned} |b(x)| &= |\sin x \cos x - 3\sin(2x)| \leq |\sin x \cos x| + |3\sin(2x)| \\ &= |\sin x| \cdot |\cos x| + 3|\sin(2x)| \leq 1 \cdot 1 + 3 \cdot 1 = 4, \quad \forall\, x \in \mathbb{R}, \end{aligned} \tag{4}$$

it follows that b is bounded on $\mathbb{R}$. We also note that

$$\lim_{x\to+\infty} \frac{1}{x^2+1} = \lim_{x\to+\infty} \frac{1}{x^2} = 0\,. \tag{5}$$

From Eq. (4) and Eq. (5), via the zero-bounded theorem (Theorem 2.3):

$$\lim_{x\to+\infty} \frac{\sin x \cos x - 3\sin(2x)}{x^2+1} = 0.$$

$\square$

Note that in Eq. (5) above we invoke a shortcut where the limit to infinity of a function defined as the ratio of two polynomials is equal to the limit of the ratio of the highest powers. Instructors who do not teach this shortcut to students should use the more detailed factorization argument instead. It is also worth noting that while this problem can be also solved with the squeeze theorem, two inequalities have to be shown, and for the factors $\sin x$, $\cos x$, $\sin(2x)$ one must choose whether to bound them from above or below. Using the zero-bounded theorem, such decision-making is not necessary, and only one inequality needs to be established.

Example 2.5 Evaluate the following limit

$$\lim_{x\to 0} x[1-\sin(1/x)].$$

Solution: Define $b(x) = 1-\sin(1/x)$, $\forall\, x \in \mathbb{R}-\{0\}$. Since

$$|b(x)| = |1-\sin(1/x)| \le 1 + |\sin(1/x)| \le 1+1 = 2, \quad \forall\, x \in \mathbb{R}-\{0\},$$

it follows that b is bounded on $\mathbb{R}-\{0\}$. We also note that $\lim_{x\to 0} x = 0$, and it follows from the zero-bounded theorem (Theorem 2.3) that

$$\lim_{x\to 0} x[1-\sin(1/x)] = 0.$$

□

In this example, one must be careful in stating the set over which the function $b(x)$ is bounded. Specifically, we note that $\mathbb{R}-\{0\}$ is a deleted neighborhood of the limit $x \to 0$ for $\delta \to +\infty$, so the statement $|b(x)| \le 2, \forall x \in \mathbb{R}-\{0\}$ satisfies the preconditions of the zero-bounded theorem. We also should emphasize that neither of the previous two examples can be evaluated via the L'Hospital theorem.

It is also possible to have zero-bounded limits appear in a more obfuscated form as in the following example:

Example 2.6 Given the function $f(x) = (1+x)^{\sin(1/x)}$ evaluate the limit $\lim_{x\to 0} f(x)$

Solution: First we note that:

$$f(x) = (1+x)^{\sin(1/x)} = \exp[\ln(1+x)\sin(1/x)], \ \forall x \in (-1,0)\cup(0,+\infty).$$

Define $b(x) = \sin(1/x)$, $\forall\, x \in (-1,0)\cup(0,+\infty)$. It follows that

$$|b(x)| = |\sin(1/x)| \le 1, \ \forall\, x \in (-1,0)\cup(0,+\infty),$$

and therefore b is bounded on $(-1,0)\cup(0,+\infty)$. We also note that

$$\lim_{x\to 0} \ln(1+x) = \ln(1+0) = \ln 1 = 0.$$

It follows from the zero-bounded theorem (Theorem 2.3) that

$$\lim_{x\to 0} \left[\ln(1+x)\sin(1/x)\right] = 0,$$

and therefore

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \exp(\ln(1+x)\sin(1/x))$$

$$= \exp(\lim_{x\to 0} \ln(1+x)\sin(1/x))$$
$$= \exp(0) = 1.$$

□

Note that the solution uses the property $a^x = \exp(x \ln a)$ and the continuity of the exponential function.

Another interesting example, brought to my attention by Dr. Lokenath Debnath, is the following limit:

Example 2.7 Given the function

$$f(x) = \frac{e^x + \sin x}{e^x + \cos x},$$

evaluate the limit $\lim_{x\to +\infty} f(x)$.

What is interesting here is that this limit mimics a L'Hospital-type problem, but it cannot be evaluated by repeated application of the L'Hospital theorem, because doing so yields a non-terminating repeating pattern:

$$\lim_{x\to+\infty} \frac{e^x + \sin x}{e^x + \cos x} = \lim_{x\to+\infty} \frac{e^x + \cos x}{e^x - \sin x} = \lim_{x\to+\infty} \frac{e^x - \sin x}{e^x - \cos x}$$
$$= \lim_{x\to+\infty} \frac{e^x - \cos x}{e^x + \sin x} = \lim_{x\to+\infty} \frac{e^x + \sin x}{e^x + \cos x} = \dots$$

By inspection it is easy to see that for large x, the e^x terms dominate both $\sin x$ and $\cos x$, thus the function becomes asymptotically equal to e^x/e^x. We therefore expect to find that $\lim_{x\to+\infty} f(x) = 1$. This suggests a solution method where the dominant e^x term is forcefully factored from both the numerator and denominator, as follows:

Solution: We note that

$$f(x) = \frac{e^x + \sin x}{e^x + \cos x} = \frac{e^x(1 + e^{-x}\sin x)}{e^x(1 + e^{-x}\cos x)} = \frac{1 + e^{-x}\sin x}{1 + e^{-x}\cos x}.$$

From the zero-bounded theorem we have

$$\begin{cases} \lim_{x\to+\infty} e^{-x} = 0 \\ |\sin x| \le 1, \quad \forall\, x \in \mathbb{R} \end{cases} \implies \lim_{x\to+\infty} e^{-x}\sin x = 0,$$

and also

$$\begin{cases} \lim_{x\to+\infty} e^{-x} = 0 \\ |\cos x| \le 1, \quad \forall\, x \in \mathbb{R} \end{cases} \implies \lim_{x\to+\infty} e^{-x}\cos x = 0.$$

It follows that

$$\lim_{x\to+\infty} f(x) = \lim_{x\to+\infty} \frac{1 + e^{-x}\sin x}{1 + e^{-x}\cos x} = \frac{1 + \lim_{x\to+\infty} e^{-x}\sin x}{1 + \lim_{x\to+\infty} e^{-x}\cos x} = \frac{1+0}{1+0} = 1.$$

□

3. Multidimensional limits

Let us now consider the case of multidimensional limits of the type encountered in multivariable Calculus. As was noted previously by Thompson and Wiggins [7], with multivariable limits it is very important to give a proper definition to ensure the existence of reasonable indeterminate forms. Thompson and Wiggins [7] surveyed an extensive collection of calculus textbooks and discovered that, in this regard, in many textbooks, the answers given to homework problems in the back of the book can be inconsistent with the limit definition adopted by the book author. For this reason, we will lay out the complete theory of multidimensional limits in some detail.

Let $f : A \to \mathbb{R}$ be a *scalar field* (i.e. mapping from vector to number) with $A \subseteq \mathbb{R}^n$ an open set and let $\mathbf{x}_0 \in \mathbb{R}^n$. We define the multidimensional limit $\mathbf{x} \to \mathbf{x}_0$ as follows: Let $N(\mathbf{x}_0, \delta)$ be an open ball with center $\mathbf{x}_0$ and radius δ, excluding the point $\mathbf{x}_0$ itself, defined as:

$$N(\mathbf{x}_0, \delta) = \{\mathbf{x} \in \mathbb{R}^n \mid 0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta\}.$$

Here $\|\cdot\|$ represents the usual Euclidean vector norm, and $\|\mathbf{x} - \mathbf{x}_0\|$ specifically is the geometric distance between $\mathbf{x}$ and $\mathbf{x}_0$. A necessary condition for defining the limit $\mathbf{x} \to \mathbf{x}_0$ is that $\mathbf{x}_0$ should be an accumulation point of the domain $A = \mathrm{dom}(f)$ of the scalar field f, otherwise, by definition, the limit does not exist. The formal definition of "accumulation point" reads:

$$\mathbf{x}_0 \text{ accumulation point of } A \Longleftrightarrow \forall \delta > 0 : N(\mathbf{x}_0, \delta) \cap A \neq \emptyset.$$

We may now write the formal definition of the multidimensional limit as follows:

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \ell \Longleftrightarrow \begin{cases} \mathbf{x}_0 \text{ accumulation point of } A \\ \forall\, \varepsilon > 0 : \exists\, \delta > 0 : \forall\, \mathbf{x} \in A \cap N(\mathbf{x}_0, \delta) : |f(\mathbf{x}) - \ell| < \varepsilon. \end{cases}$$

Note that, using the terminology of Thompson and Wiggins [7], this is a type-GF definition.

The standard argument in most textbooks, at this point, is to establish the basic limit properties, the continuity of polynomial functions, and the squeeze theorem. Given all that, we now add to the above the zero-bounded theorem, stated as follows:

Definition 3.1: Let $f : A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ be a given scalar field and let $B \subseteq A$. Then,

$$f \text{ bounded on } B \Longleftrightarrow \exists\, b > 0 : \forall\, \mathbf{x} \in B : |f(\mathbf{x})| \leq b.$$

Theorem 3.2: *Let $b : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ be two given scalar fields. Let $\mathbf{x}_0 \in \mathbb{R}^n$ be an accumulation point of A. Then,*

$$\begin{cases} \exists\, \delta > 0 : b \text{ bounded on } N(\mathbf{x}_0, \delta) \\ \lim_{\mathbf{x} \to \mathbf{x}_0} g(\mathbf{x}) = 0 \end{cases} \Longrightarrow \lim_{\mathbf{x} \to \mathbf{x}_0} [b(\mathbf{x}) g(\mathbf{x})] = 0.$$

This is essentially the same theorem as for the single-variable case. The main difference is that in the multidimensional case, the concept of the single-variable generalized neighborhood is replaced by the open ball $N(\mathbf{x}_0, \delta)$. Otherwise, the above definitions, theorems, and proofs are virtually identical to those corresponding to the one-dimensional case.

As it turns out, in most textbooks, in the given homework problems for multidimensional limits, the corresponding limit will either not exist, or is an indeterminate 0/0 form that can be resolved by some cancellation, or is a zero-bounded limit. Below is an example of a zero-bounded limit:

Example 3.3 Evaluate the following limit

$$\lim_{(x,y)\to(0,0)} \frac{3x^2(x^2-y^2)}{x^2+y^2}.$$

Solution: Define $b(x,y) = \frac{3x^2}{x^2+y^2}$, $\forall (x,y) \in \mathbb{R}^2 - \{(0,0)\}$. It follows that

$$\begin{aligned} |b(x,y)| &= \left|\frac{3x^2}{x^2+y^2}\right| = \frac{3x^2}{x^2+y^2} \le \frac{3x^2+3y^2}{x^2+y^2} \\ &= \frac{3(x^2+y^2)}{x^2+y^2} = 3, \quad \forall\ (x,y) \in \mathbb{R}^2 - \{(0,0)\}, \end{aligned}$$

and therefore b is bounded on $\mathbb{R}^2 - \{(0,0)\}$. By polynomial continuity, we also have

$$\lim_{(x,y)\to(0,0)} (x^2-y^2) = 0^2 - 0^2 = 0,$$

and it follows from the zero-bounded theorem that

$$\lim_{(x,y)\to(0,0)} \frac{3x^2(x^2-y^2)}{x^2+y^2} = 0.$$

□

For the case of limits that do not exist, it is usually sufficient to simply evaluate them across two different paths and show that different answers are obtained. Although most textbooks tend to explain this reasonably well, most expositions can be improved by introducing more formal notation as follows. Let $\mathbf{x}_0 \in \mathbb{R}^n$ be a given point, and let $\mathcal{P}(\mathbf{x}_0|A)$ be the set of all continuous mappings $\gamma : (0,a) \to A - \{\mathbf{x}_0\}$ with $a \in (0,+\infty)$ and $A \subseteq \mathbb{R}^n$ such that:

$$\lim_{t\to 0^+} \|\gamma(t) - \mathbf{x}_0\| = 0.$$

Then, for any path $\gamma \in \mathcal{P}(\mathbf{x}_0|A)$ we define the path-restricted limit as:

$$\lim_{\gamma} f = \lim_{t\to 0^+} f(\gamma(t)).$$

To show that the limit of $f(\mathbf{x})$ with $\mathbf{x} \to \mathbf{x}_0$ does not exist, it is sufficient to define two paths $\gamma_1, \gamma_2 \in \mathcal{P}(\mathbf{x}_0|A)$ such that $\lim_{\gamma_1} f \neq \lim_{\gamma_2} f$. In more formal terms, the methodology for showing the non-existence of a multidimensional limit amounts to invoking the following statement:

$$(\exists\ \gamma_1, \gamma_2 \in \mathcal{P}(\mathbf{x}_0|A) : \lim_{\gamma_1} f \neq \lim_{\gamma_2} f) \Longrightarrow \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) \text{ does not exist.} \qquad (6)$$

On the other hand, Dobbs [8] showed that if all path limits $\lim_\gamma f$ agree over all paths $\gamma \in \mathcal{P}(\mathbf{x}_0|A)$, then the multidimensional limit $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ will exist and agree with them. The corresponding statement reads:

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \ell \Longleftrightarrow \forall\gamma \in \mathcal{P}(\mathbf{x}_0|A) : \lim_{\gamma} f = \ell. \tag{7}$$

As is correctly noted in most textbooks, it is possible to have multidimensional limits in which all path-restricted limits employing *linear* paths may agree with each other but disagree with a path-restricted limit along a nonlinear path. A classic example, cited by both Stewart [2, 5] and Apostol [6], rewritten in terms of our notation, is the following:

Example 3.4 Consider the function

$$f(x,y) = \frac{xy^2}{x^2+y^4}, \quad \forall\ (x,y) \in \mathbb{R}^2 - \{(0,0)\},$$

and let $\gamma(\theta)$ be a linear path towards $(0,0)$ defined as

$$\gamma(\theta) : \begin{cases} x = t\cos\theta \\ y = t\sin\theta \end{cases}, \quad \text{with } t \to 0^+.$$

Then, it can be shown that

$$\forall\ \theta \in [0, 2\pi] : \lim_{\gamma(\theta)} f = 0,$$

and yet $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Solution: Let $\theta \in [0, 2\pi]$ be given. We note that

$$\begin{aligned}\lim_{\gamma(\theta)} f &= \lim_{\gamma(\theta)} \frac{xy^2}{x^2+y^4} = \lim_{t\to 0^+} \frac{(t\cos\theta)(t\sin\theta)^2}{(t\cos\theta)^2 + (t\sin\theta)^4} \\ &= \lim_{t\to 0^+} \frac{t^3\cos\theta\sin^2\theta}{t^2\cos^2\theta + t^4\sin^4\theta} = \lim_{t\to 0^+} \frac{t\cos\theta\sin^2\theta}{\cos^2\theta + t^2\sin^4\theta}.\end{aligned}$$

Let us distinguish between the following two cases:
(a) Assume that $\cos\theta \neq 0$. Then:

$$\lim_{\gamma(\theta)} f = \frac{0\cos\theta\sin^2\theta}{\cos^2\theta + 0^2\sin^4\theta} = \frac{0}{\cos^2\theta + 0} = 0.$$

Note that the $\cos\theta \neq 0$ assumption is needed to avoid a possible $0/0$ indeterminate form.
(b) Assume that $\cos\theta = 0$. It follows that either $\sin\theta = 1$ or $\sin\theta = -1$, and therefore

$$\lim_{\gamma(\theta)} f = \lim_{t\to 0^+} \frac{t\cdot 0\cdot(\pm 1)^2}{0^2 + t(\pm 1)^4} = \lim_{t\to 0^+} \left(\frac{0}{t^2}\right) = 0.$$

Putting both cases together, we have shown that $\forall\ \theta \in [0, 2\pi] : \lim_{\gamma(\theta)} f = 0$.

Now consider the path γ_0 defined as:

$$\gamma_0 : \begin{cases} x = t^2 \\ y = t \end{cases} , \quad \text{with } t \to 0^+.$$

Then, the path-restricted limit of f along γ_0, reads:

$$\begin{aligned}\lim_{\gamma_0} f = \lim_{\gamma_0} \frac{xy^2}{x^2+y^4} &= \lim_{t\to 0^+} \frac{t^2 t^2}{(t^2)^2 + t^4} \\ &= \lim_{t\to 0^+} \frac{t^4}{t^4+t^4} = \lim_{t\to 0^+} \frac{t^4}{2t^4} = \frac{1}{2}.\end{aligned}$$

It follows that $\lim_{\gamma_0} f \neq \lim_{\gamma(\theta)} f$ for any $\theta \in [0, 2\pi]$ and therefore the limit $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist. □

4. Multidimensional limits and continuity

In surveying various textbooks with respect to their discussion of multidimensional limits, I have noticed problematic explanations that can easily lead to serious conceptual misunderstandings. For an egregious example, let us consider the following quotation from pp. 631-632 of Varburg et al. [3]:

> "It is often easier to analyze limits of functions of two variables, especially limits at the origin, by changing to polar coordinates. The important point is that $(x, y) \to (0, 0)$ if and only if $r = \sqrt{x^2 + y^2} \to 0$. Thus, limits for functions of two variables can sometimes be expressed as limits involving just one variable r.
> **Example 3:** Evaluate the following limits, if they exist:
>
> $$\text{(a)} \quad \lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{3x^2+3y^2} \qquad \text{(b)} \quad \lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$$
>
> (a) Changing to polar coordinates and using L'Hospital's Rule, we have
>
> $$\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{3x^2+3y^2} = \lim_{r\to 0} \frac{\sin(r^2)}{3r^2} = \frac{1}{3} \lim_{r\to 0} \frac{2r\cos r^2}{2r} = \frac{1}{3}.$$
>
> (b) Again, changing to polar coordinates gives
>
> $$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{r\to 0} \frac{r\cos\theta\, r\sin\theta}{r^2} = \cos\theta\sin\theta.$$
>
> Since this limit depends on θ, straight line paths to the origin will lead to different limits. Thus, this limit does not exist."

While the solution to the second subquestion is essentially correct, the solution to the first subquestion seems to be justified on a false premise; changing to polar coordinates and taking the limit $r \to 0$ seems to suggest that it is sufficient to consider path-restricted limits over all *linear* paths approaching the limit point and show only that all linear paths are in agreement. This idea is then inadvertantly reinforced by the comments after the second subquestion as well as by the vague and poorly-phrased statement that "$(x, y) \to (0, 0)$ if and only if $r \to 0$".

As shown by the previous counterexample, agreement between the path-restricted limits over all linear paths is *not* sufficient for establishing the existence of the multidimensional limit itself. On the other hand, we will now argue that the authors' argument can still be salvaged if it is justified via continuity and the composition theorem. While both topics are discussed by the authors, the previously quoted example appears *before* their discussion of the composition theorem and no connection is made later between the composition theorem and this "change to polar coordinates" technique.

We will now state the definition of continuity and the composition theorem, and then use the composition theorem (Theorem 4.2) to provide a mathematically correct solution for the limit in the example quoted above. Unfortunately, a correct and rigorous solution that directly uses the composition theorem becomes more complicated and less student friendly. This is probably the underlying reason for tolerating non-rigorous explanations, as in the above-quoted textbook excerpt. On the other hand, if we use instead a corollary (Proposition 4.3) of the composition theorem, we can have a more pedagogical solution which is still completely rigorous. The price to be paid though is that Proposition 4.3 has a somewhat weaker range of applicability than the original composition theorem. For this reason, in the example below, we provide two solutions for the same problem: a solution using the composition theorem directly, and a solution based on Proposition 4.3. Both solutions are equally rigorous, although the second solution is more pedagogical than the first. We also use Eq. (7) by Dobbs [8], to show a very simple proof of the composition theorem for the multi-dimensional case by piggybacking on the one-dimensional composition theorem, which we presume has been proven earlier in single-variable calculus.

Definition 4.1: Let $f : A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ be a scalar field, and let $x_0 \in A$ be given. Also let $B \subseteq A$ be given. We say that

(1) f continuous at $\mathbf{x}_0 \Longleftrightarrow \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$.
(2) f continuous on $B \Longleftrightarrow \forall\, \mathbf{x}_0 \in B : f$ continuous at $\mathbf{x}_0$.

Theorem 4.2: *Let $f : A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ be a scalar field, let $g : B \to \mathbb{R}$ with $B \subseteq \mathbb{R}$ be a function, and let $\mathbf{x}_0 \in A$. Then,*

$$\left.\begin{array}{l} f \text{ continuous at } \mathbf{x}_0 \\ g \text{ continuous at } f(\mathbf{x}_0) \end{array}\right\} \Longrightarrow h = g \circ f \text{ continuous at } \mathbf{x}_0.$$

Proof: Let $C = \operatorname{dom}(g \circ f)$ and let $\gamma \in \mathcal{P}(\mathbf{x}_0|C)$ be given. Since $C \subseteq A$, it follows that $\gamma \in \mathcal{P}(\mathbf{x}_0|A)$ as well. We may therefore define

$$\varphi(t) = \begin{cases} f(\gamma(t)), & \text{if } t > 0 \\ f(\mathbf{x}_0), & \text{if } t = 0. \end{cases}$$

Since f is continuous in $\mathbf{x}_0$, it follows that

$$\lim_{t\to 0^+} \varphi(t) = \lim_{t\to 0^+} f(\gamma(t)) = \lim_{\gamma} f = \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0) = \varphi(0)$$
$$\Longrightarrow \varphi \text{ continuous at } t = 0.$$

Note that the existence and evaluation of the limit $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ follows from the assumption that f is continuous in $\mathbf{x}_0$. The path limit $\lim_{\gamma} f$ is thus restricted to being equal to $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$. From the one-dimensional composition theorem, we obtain

$$\left.\begin{array}{l}\varphi \text{ continuous at } t = 0\\ g \text{ continuous at } \varphi(0)\end{array}\right\} \Longrightarrow \lim_{t\to 0^+} g(\varphi(t)) = g(\varphi(0)),$$

and it follows that

$$\begin{aligned}\lim_{\gamma} h &= \lim_{t\to 0^+} h(\gamma(t)) = \lim_{t\to 0^+} g(f(\gamma(t))) = \lim_{t\to 0^+} g(\varphi(t)) = g(\varphi(0))\\ &= g(f(\mathbf{x}_0)) = h(\mathbf{x}_0).\end{aligned}$$

We have thus shown that

$$\begin{aligned}(\forall \gamma \in \mathcal{P}(\mathbf{x}_0|C) : \lim_{\gamma} h = h(\mathbf{x}_0)) &\Longrightarrow \lim_{\mathbf{x}\to\mathbf{x}_0} h(\mathbf{x}) = h(\mathbf{x}_0)\\ &\Longrightarrow h \text{ continuous at } \mathbf{x} = \mathbf{x}_0\end{aligned}$$

□

Proposition 4.3: *Let $f : A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ be a scalar field, let $g : B \to \mathbb{R}$ with $B \subseteq \mathbb{R}$ be a function, and let $\mathbf{x}_0 \in \mathbb{R}^n$ be an accumulation point of A and ℓ_0 be an accumulation point of B. Then,*

$$\left.\begin{array}{l}\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \ell_0\\ \lim_{t\to\ell_0} g(t) = \ell\\ \exists\, \delta > 0 : \forall\, \mathbf{x} \in N(\mathbf{x}_0, \delta) : f(\mathbf{x}) \neq \ell_0\end{array}\right\} \Longrightarrow \lim_{\mathbf{x}\to\mathbf{x}_0} g(f(\mathbf{x})) = \ell.$$

Proof: We define the following functions:

$$F(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \text{if } \mathbf{x} \in A - \{\mathbf{x}_0\}\\ \ell_0, & \text{if } \mathbf{x} = \mathbf{x}_0,\end{cases}$$

$$G(t) = \begin{cases} g(t), & \text{if } t \in B - \{\ell_0\}\\ \ell, & \text{if } t = \ell_0.\end{cases}$$

Then, by definition, F is continuous at $\mathbf{x} = \mathbf{x}_0$ and G is continuous at $t = \ell_0 = F(\mathbf{x}_0)$. From the composition theorem, it follows that $H = G \circ F$ is continuous at $\mathbf{x} = \mathbf{x}_0$, and consequently

$$\lim_{\mathbf{x}\to\mathbf{x}_0} G(F(\mathbf{x})) = \lim_{\mathbf{x}\to\mathbf{x}_0} H(\mathbf{x}) = H(\mathbf{x}_0) = G(F(\mathbf{x}_0)) = G(\ell_0) = \ell.$$

Unfortunately, we cannot deduce that $\lim_{\mathbf{x}\to\mathbf{x}_0} g(f(\mathbf{x})) = \ell$ unless we can invoke the hypothesis $\exists\, \delta > 0 : \forall\, \mathbf{x} \in N(\mathbf{x}_0, \delta) : f(\mathbf{x}) \neq \ell_0$ as follows: Given the appropriate $\delta > 0$,

$$\begin{aligned}\forall \mathbf{x} \in N(\mathbf{x}_0, \delta) : G(F(\mathbf{x})) &= G(f(\mathbf{x})) && [\text{because } \mathbf{x} \neq \mathbf{x}_0]\\ &= g(f(\mathbf{x})). && [\text{because by hypothesis, } f(\mathbf{x}) \neq \ell_0]\end{aligned}$$

The assumption $\mathbf{x} \neq \mathbf{x}_0$ follows from $\mathbf{x} \in N(\mathbf{x}_0, \delta)$ and $f(\mathbf{x}) \neq \ell_0$ is given by hypothesis. It follows that

$$\lim_{\mathbf{x}\to\mathbf{x}_0} g(f(\mathbf{x})) = \lim_{\mathbf{x}\to\mathbf{x}_0} G(F(\mathbf{x})) = \ell.$$

□

Example 4.4 Given the function

$$f(x,y) = \frac{\sin(x^2+y^2)}{3x^2+3y^2},$$

evaluate the limit $\lim_{(x,y)\to(0,0)} f(x,y)$.

Solution: Define the auxiliary functions

$$g(x) = \begin{cases} \sin(x)/(3x), & \text{if } x \in \mathbb{R} - \{0\} \\ 1/3, & \text{if } x = 0, \end{cases}$$
$$h(x,y) = x^2 + y^2,\ \forall\ (x,y) \in \mathbb{R}^2.$$

Then, by definition,

$$\forall\ (x,y) \in \mathbb{R}^2 - \{(0,0)\} : f(x,y) = g(h(x,y)).$$

We note that at $(x,y) = (0,0)$,

$$\lim_{(x,y)\to(0,0)} h(x,y) = \lim_{(x,y)\to(0,0)} (x^2+y^2) = 0^2 + 0^2 = 0 = h(0,0)$$
$$\Longrightarrow h \text{ continuous at } (x,y) = (0,0),$$

and at $x = h(0,0) = 0$ we have

$$\lim_{x\to 0} g(x) = \lim_{x\to 0} \frac{\sin x}{3x} = \frac{1}{3} = g(0)$$
$$\Longrightarrow g \text{ continuous at } x = h(0,0).$$

It follows that

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} g(h(x,y)) = g(h(0,0)) = g(0) = 1/3.$$

□

Note that the application of the composition theorem on the solution above requires an explicit definition of $g(x)$ and $h(x,y)$ to maintain rigor. Using Proposition 4.3, we can eliminate the need to explicitly define $g(x)$ and $h(x,y)$ to obtain the following simpler solution:

Solution: We note that

$$\lim_{(x,y)\to(0,0)} (x^2+y^2) = 0^2 + 0^2 = 0, \tag{8}$$

and for $t = x^2 + y^2$,

$$\lim_{t\to 0} \frac{\sin t}{3t} = \frac{1}{3}. \tag{9}$$

We also note that

$$x^2 + y^2 \neq 0, \forall (x, y) \in \mathbb{R}^2 - \{(0, 0)\}. \tag{10}$$

It follows that

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^2 + y^2)}{3x^2 + 3y^2} = \frac{1}{3}. \tag{11}$$

□

It is important to note that the statement $x^2 + y^2 \neq 0, \forall (x, y) \in \mathbb{R} - \{(0, 0)\}$ is needed to make the solution mathematically rigorous. That is the price that has to be paid for circumventing the direct application of the composition theorem, as in the previous solution.

5. Conclusion

As Thompson and Wiggins previously noted [7], the treatment of limits in calculus courses tends to be careless. The carelessness is often rationalized under the excuse that calculus courses do not need to be rigorous. On the other hand, whereas we may not be going to be proving the theorems themselves, the fact remains that even the arguments needed to evaluate limits are themselves proofs, and as such, they have to be done rigorously, if we wish the arguments to be mathematically correct. Pedagogy should not be employed as an excuse for teaching solution techniques in an incomplete or non-rigorous manner, thereby reinforcing serious misconceptions.

For problems involving zero-bounded limits I have suggested a simpler and yet rigorous method of evaluation as an alternative to the squeeze theorem. The method reinforces skills needed in more advanced courses, and it is useful for both single-variable and especially for multidimensional limits. We have also highlighted the role of continuity and the composition theorem in the evaluation of multi-dimensional limits. It should not be difficult for instructors to incorporate these techniques in their calculus courses.

Acknowledgements

The teaching of the zero-bounded limit theorem used to be part of the national 12^{th} grade high-school curriculum in Greece. I wish to thank Dr. Lokenath Debnath for bringing the rather amusing limit of Example 2.7 to my attention. I am also indebted to an anonymous referee for a detailed and meticulous review of the paper and for bringing the paper by Dobbs [8] to my attention. The reproduction of a short excerpt from Ref. [3] is permitted under the fair dealing provision of UK Copyright law, for the purposes of criticism and review, occuring in the paragraphs immediately following the quoted excerpt.

Appendix A. Squeeze theorem

In this appendix we prove the squeeze theorem as a consequence of Lemma 2.1.

Theorem A.1: *Let $f : A \to \mathbb{R}$, $g_1 : A \to \mathbb{R}$, and $g_2 : A \to \mathbb{R}$ be three functions and let σ be an accumulation point of A. Then,*

$$\left.\begin{array}{l}\forall\ x \in N(\sigma,\delta) : g_1(x) \leq f(x) \leq g_2(x)\\ \lim\limits_{x\to\sigma} g_1(x) = \lim\limits_{x\to\sigma} g_2(x) = \ell \in \mathbb{R}\end{array}\right\} \Longrightarrow \lim_{x\to\sigma} f(x) = \ell.$$

Proof: Let $x \in N(\sigma,\delta)$ be given. Then

$$\begin{aligned}&g_1(x) \leq f(x) \leq g_2(x) \Rightarrow 0 \leq f(x) - g_1(x) \leq g_2(x) - g_1(x)\\ &\Rightarrow 0 \leq |f(x) - g_1(x)| \leq g_2(x) - g_1(x),\end{aligned}$$

and therefore

$$\forall\ x \in N(\sigma,\delta) : |f(x) - g_1(x)| \leq g_2(x) - g_1(x).$$

We also note that

$$\lim_{x\to\sigma}(g_2(x) - g_1(x)) = \lim_{x\to\sigma} g_2(x) - \lim_{x\to\sigma} g_1(x) = \ell - \ell = 0,$$

so the assumptions of Lemma 2.1 are satisfied, consequently,

$$\lim_{x\to\sigma}(f(x) - g_1(x)) = 0,$$

and it follows that

$$\begin{aligned}\lim_{x\to\sigma} f(x) &= \lim_{x\to\sigma}(f(x) - g_1(x) + g_1(x))\\ &= \lim_{x\to\sigma}(f(x) - g_1(x)) + \lim_{x\to\sigma} g_1(x)\\ &= 0 + \ell = \ell.\end{aligned}$$

□

References

[1] J. Stewart *Single Variable Calculus: Early Transcendentals*, Brooks/Cole Publishing Company, Pacific Grove CA, 1999.
[2] J. Stewart *Calculus*, Thomson Learning, Belmont CA, 2003.
[3] D. Varberg, E. Purcell, and S. Rigdon *Calculus*, Pearson Prentice Hall, New Jersey NJ, 2007.
[4] J. Haas, M. Weir, and G. Thomas *University Calculus. Early Transcendentals*, Pearson Education, Boston MA, 2012.
[5] J. Stewart *Essential Calculus. Early Transcedentals*, Thomson Brooks/Cole, Belmont CA, 2007.
[6] T. Apostol *Calculus. I*, John Wiley & Sons, New York NY, 1967.
[7] T. Thompson and K. Wiggins, *The treatment of limits in the calculus: are we becoming careless?*, Int. J. Math. Ed. Sci. Tech. 21 (1990), pp. 235–242.
[8] D. Dobbs, *Limits along curves determine limits*, Int. J. Math. Ed. Sci. Tech. 28 (1997), pp. 303–307.