Brouwer's ϵ-fixed point from Sperner's lemma[1]

Dirk van Dalen

It is by now common knowledge that Brouwer gave mathematics in 1911 a miraculous tool, the fixed point theorem, and that later in life, he disavowed it. It usually came as a shock when he replied to the question "is the fixed point theorem correct ?" with a point blank "no". This rhetoric exchange deserves some elucidation. At the time that Brouwer did his revolutionary topological work, he has suspended his constructive convictions for the time being. He was well aware that he was using the principle of the excluded middle, indeed in [Brouwer 1919], p. 950, he remarked that "In my philosophy-free mathematical papers I have regularly used the old methods, while at the same time attempting to deduce only those results, of which I could hope that they would find a place and be of value, if necessary in a modified form, in the new doctrine after the carrying out of a systematic construction of intuitionistic set theory". And in the case of the fixed point theorem we are presented with exactly such a result. From the intuitionistic point of view the theorem is not correct because the fixed point that is promised can in general not be found, that is to say, approximated.

A counterexample to the fixed point theorem.

First let us recall that from a constructive point of view the equality between real numbers is undecidable, in the sense that we have no means to decide for any two reals a and b whether $a = b$ or $a \neq b$. The same holds for the order relation on the reals : it is undecidable if $a < b$ or $a = b$ or $a > b$ − in logical formulation we have no grounds for asserting $a < b \vee a = b \vee a > b$. Here $\vee$ (or) is interpreted in its constructive, strong meaning : $A \vee B$ holds if we can (i) pick one of A and B and (ii) provide a proof for it. Existence statements have a similar constructive interpretation : $\exists x A(x)$ holds if we can (i) construct an object (say a natural number) n, and (ii) prove $A(n)$. A proof in the constructive context is also a construction, so construction is here the key notion. The specific notion "construction" is left open, because in principle it depends on what we are talking about, see for the proof interpretation [Troelstra, A.S. and D. van Dalen 1988], section 3, [Dalen, D. van 2008] , [Soerensen, M.H. and P. Urzyczyn 2006], There are in fact helpful formal systems that capture the practice of proving, see [Dalen, D. van 2008] ch. 5. Whereas it is intuitively clear what we mean by a construction operating on natural numbers (finite, discrete objects), it all becomes more complicated when dealing with infinite objects, like real numbers. Here the key notion is "approximation".

Brouwer designed a technique for demonstrating the undecidability of certain facts from classical mathematics. It produces so-called *Brouwerian counte-*

[1] Update 01/14/2010. I am indebted to Daniel Méhkeri for spotting an oversight. A first version was presented at the International Conference Systemics, Cybernetics, and Informatics, Hyderabad 2009. *How to constructivize Brouwer's fixed point theorem*, Proceedings, Pentagram 2009, 654–656.

rexamples. The best known form is the one using the decimal expansion of π : we compute simultaneously the decimal expansion of π and a Cauchy sequence to be specified. We use $N(k)$ as an abbreviation for 'the decimals $p_{k-89}, \dots, p_k$ of π are all 9'. Now we define

$$a_n = \begin{cases} (-2)^{-n} & \text{if } \forall k \leq n \neg N(k) \\ (-2)^{-k} & \text{if } k \leq n \text{ and } N(k) \text{ and } \forall \ell < k \neg N(\ell) \end{cases}$$

a_n starts as an oscillating sequence of negative powers of -2. Should we hit upon a sequence of 90 nines in the expansion of π, a_n becomes constant from there on :

$1, -\frac{1}{2}, \frac{1}{4}, -\frac{1}{8}, \dots, (-2)^{-k}, (-2)^{-k}, (-2)^{-k}, \dots$

The sequence a_n satisfies the Cauchy condition and in that sense determines a real number a. The sequence is well-defined, and we can, in principle, check $N(n)$ for each n. For this a we have

$$\begin{aligned} &a > 0 \Leftrightarrow N(k) \text{ holds the first time for an even number,} \\ &a < 0 \Leftrightarrow N(k) \text{ holds the first time for an odd number,} \\ &a = 0 \Leftrightarrow N(k) \text{ holds for no } k, \end{aligned}$$

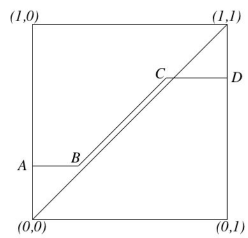
and thus also

$$\begin{aligned} &a \geq 0 \Leftrightarrow \neg a < 0 \Leftrightarrow N(k) \text{ does not hold the first time at an odd number,} \\ &a \leq 0 \Leftrightarrow \neg a > 0 \Leftrightarrow N(k) \text{ does not hold the first time at an even number,} \end{aligned}$$

where $x \leq y$ is defined as $\neg y < x$.

Since we have no means to determine the existence of a number k with $N(k)$, none of the above cases can be asserted. E.g. in order to assert $a > 0$ we have to construct (in finitely many steps) an even number k such that the digit p_{k+1} of π is the first one preceded by ninety nines. In the absence of such a construction we are not allowed to assert $a > 0$.

We will proceed to a Brouwerian counter example to the fixed point theorem. We consider the one dimensional case. Let $f : [0,1] \to [0,1]$ be a piecewise linear function its graph passing through the points $A = (0, \frac{1}{4}+a), B = (\frac{1}{4}, \frac{1}{4}+a), C = (\frac{3}{4}, \frac{3}{4}+a), D = (1, \frac{3}{4}+a)$, where a is one of those points for which it is unknown if $a > 0, a = 0, a < 0$.



A fixed point of f is obtained by intersecting the graph of f with the diagonal. Say there is a fixed point u.

One of the basic properties of the ordering on $\mathbb{R}$ is

$$x < y \rightarrow x < z \vee z < y,$$

hence $\frac{1}{4} < u$ or $u < \frac{3}{4}$. In the first case we have $a \leq 0$ and in the second case $a \geq 0$. Hence the fixed point theorem tells us that we can decide $a \leq 0 \vee a \geq 0$. Quod non. Conclusion : the fixed point theorem is not constructively correct.

Constructivizing the fixed point theorem

In his Berlin lectures Brouwer pointed out that there are constructive substitutes for certain classical theorems, including the fixed point theorem [Brouwer 1992], p. 57. Another version was treated in [Brouwer 1952].

In the present paper we are going to show that on the basis of Sperner's lemma, a constructivization of the fixed point can be given. We will deal with the two-dimensional case, and we will consider the fixed point theorem on a triangle. Thus we can use the technique of Sperner's lemma.

Consider a triangle ΔABC. We triangulate Δ, i.e. Δ is partitioned into small triangles such that adjacent triangles share sides. The vertices of the triangles are labeled with the numbers 0,1,2 by a function f. We fix $f(A) = 0, f(B) = 1, f(C) = 2$, and furthermore the points on one of the sides of Δ share their labels with the endpoints : if P is on AB, then $f(P) = f(A)$ or $f(P) = f(B)$, similarly for BC and CA. An element of the triangulation is called "full" if its vertices have distinct labels.

Sperner's lemma : at least one element of the triangulation is full.

To make sure that our results are constructively correct, let us go over the standard proof of the lemma.

We introduce a graph for the triangulation that has as its nodes the elements of the triangulation, plus one extra node (think of putting a dot in each triangle, and one dot outside Δ). We now define the edges of the graph : two nodes are connected by an edge if they share a side labeled 01.

We mention an auxiliary lemma :

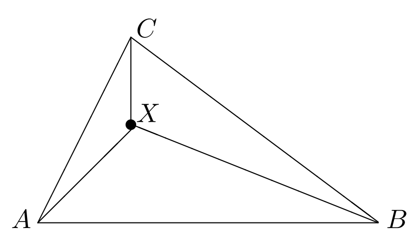
lemma *A finite graph $\mathcal{G}$ contains an even number of odd nodes*,

where a node is called *even* if it is the endpoint of an even number of edges. A node that is not even is called *odd*. The proof is an easy induction.

We now continue the proof of Sperner's lemma. On the line AB there are only 0- and 1-points. A simple argument shows that there are an odd number of adjacent 01 sides, hence the 'outside' node is odd. And therefore the preceding lemma tells us that there must be an odd node inside. One immediately checks that a triangle that stands for an odd node of the graph, has to be full. □

We will now use Sperner's lemma to establish Brouwer's fixed point theorem in the classical setting. This will show where the proof by contradiction comes in.

Before we do so, we will introduce some fact about barycentric coordinates. These allow us to give a streamlined proof of the fixed point theorem.
Consider a triangle, say our Δ. At the vertices of the triangle we deposit certain weights, say w_A, w_B, w_C. This determines uniquely the centre of gravity. The numbers w_A, w_B, w_C are called the barycentric coordinates of this centre ; they are homogeneous in the sense that $\lambda w_A, \lambda w_B, \lambda w_C$ determines exactly the same point. Thus one can normalize the coordinates such that $w_A + w_B + w_C = 1$. There is a convenient characterization of the normalized coordinates :



The barycentric coordinates are exactly the areas of the triangles formed by X and the vertices of Δ (see [Henle 1979]).

Brouwer's classical fixed point theorem

If $f : \Delta \to \Delta$ is continuous, then f has a fixed point.

Proof. Assume that f has no fixed point. We consider the class of triangulations with arbitrarily small diameters. Let a triangulation of Δ be given, we assign labels to the vertices of the triangulation triangles as follows. For a point x we determine the image $f(x)$. Since we have $x_0+x_1+x_2 = f(x)_0+f(x)_1+f(x)_2 = 1$, and since f has no fixed point, we get $f(x)_i < x_i$ for at least one i. Label x with such an i. By looking at the areas of the triangles belonging to x and $f(x)$ we see that the condition of the points on the sides of Δ are satisfied. Hence, by Sperner's lemma, there is a full triangle in each triangulation. Since Δ is compact there is limit point z of the collection of all full triangles. I.e. we have a converging sequence of full triangles of descending size. In each of these triangles we have the inequalities $f(u)_i < u_i, f(v)_j < v_j, f(w)_k < w_k$, for the vertices u, v, w and the appropriate i, j, k. Hence for the limit point z we have $f(z)_0 \leq z_0, f(z)_1 \leq z_1, f(z)_2 \leq z_2$; combined with $f(z)_0 + f(z)_1 + f(z)_2 = z_0 + z_1 + z_2 = 1$, this yields $f(z)_0 = z_0, f(z)_1 = z_1, f(z)_2 = z_2$. That is, z is a fixed point – contradiction, hence f has a fixed point. □

This indirect proof uses the principle of the excluded middle, and moreover, the Brouwerian counterexample tells us that we cannot do better.

How to repair the fixed point theorem, or find a constructive substitute for it ? Brouwer noted that although we cannot find a point that remains exactly in its place, we may look for points that move very little. In precise terms : can we find for each ϵ a point x such that $|x - f(x)| < \epsilon$? We will call such a point an ϵ-fixed point.

In the proof of the ϵ-fixed point theorem for the triangle will use the same technique as in the classical case, but we have to refine the labeling technique. We consider an equilateral triangle Δ with sides of length 1. For each n we introduce a triangulation, $\mathcal{T}_n$, obtained by dividing the sides in n equal segments of length 2^{-n}, and drawing lines parallel to the sides of Δ. As we are looking for an ϵ-fixed point the exact nature of the triangulation is not so important ; the main thing is to get arbitrarily fine $\mathcal{T}_n$'s.

There is one more point that we should mention before proceeding : in ordinary classical mathematics continuity implies uniform continuity ; in constructive mathematics this is in general not the case. In recursive mathematics there are continuous functions on [0,1] that are not uniformly continuous, while in Brouwer's universe all functions on [0,1] are automatically uniformly continuous. So we will have to stipulate explicitly that we are dealing with uniformly continuous functions on Δ.

We choose a specific formulation of uniform continuity, this will simplify the proof. Consider the triangulation $\mathcal{T}_k$. Here is the standard formulation : $\forall\epsilon\exists\delta(d(x, y) < \delta \rightarrow d(f(x), f(y)) < \epsilon$; here we take $\epsilon = 2^{-n}$ and $\delta = 2^{-k}+2^{-k-2}$ (this means that the formulation actually reads as $\forall n\exists k \ldots$). We abbreviate δ as $2^{-k} + 2\tau$. Obviously we may take $\delta < \epsilon$.

Now we proceed to get a labeling. We cannot just mimic the previous one, because the ordering relations are undecidable.

By $x_0 < x_0 + \tau$, we get $x_0 < f(x)_0$ or $f(x)_0 < x_0 + \tau$. If $x_0 < f(x)_0$, then it follows from $x_0 + x_1 + x_2 = f(x)_0 + f(x)_1 + f(x)_2 = 1$ that $x_1 + x_2 > f(x)_1 + f(x)_2$. It is intuitionistically correct to conclude now that $x_1 > f(x)_1$ or $x_2 > f(x)_2$. Therefore we have for every x : $f(x)_0 < x_0[+\tau]$ or $f(x)_1 < x_1[+\tau]$ or $f(x)_2 < x_2[+\tau]$. Hence we may label x with an i such that $f(x)_i < x_i + \tau$ or $f(x)_i < x_i$.

We now have to check the conditions on the vertices on the sides of Δ. Let x be situated on the line through 0 and 1. By the above outlined method we get $f(x)_i < x_i + \tau$ for some i. The case to be considered is $i = 2$; it cannot be ruled out a priori, so we have to get around it in some way. Assume therefore that we have a label 2 because $f(x)_2 < x_2 + \tau$. Since $x_2 = 0$, we get $f(x)_2 < \tau$.

$x_0 + x_1 + x_2 = f(x)_0 + f(x)_1 + f(x)_2$ and
$x_2 = 0, f(x)_2 < \tau$.

Hence $x_0 + x_1 < f(x)_0 + f(x)_1 + \tau$. It now follows that

$$(x_0 < f(x)_0) \vee (x_1 < f(x)_1 + \tau)$$

In the first case we obtain, as before, $f(x)_1 < x_1$. So let us now consider the second case.

Let us put $u = \tau - f(x)_2$. We note that $u > 0$ and $\tau > u$;

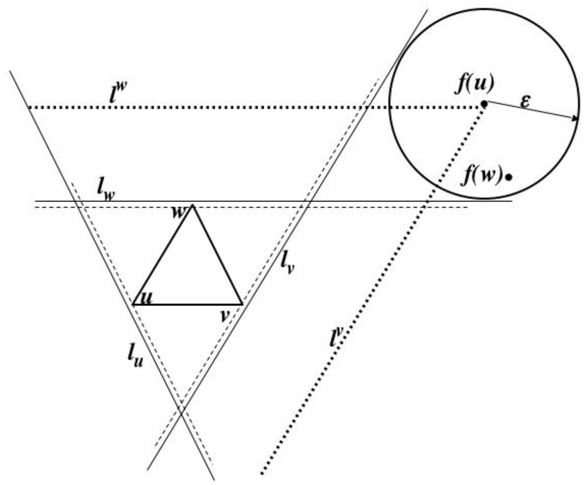
now $x_0 + x_1 + \tau = f(x)_0 + f(x)_1 + f(x)_2 + \tau$ and $x_1 < f(x)_1 + \tau$, hence $\tau + x_0 > f(x)_0 + f(x)_2$. So $x_0 > f(x)_0 + f(x)_2 - \tau = f(x)_0 - u$, and $f(x)_0 < f(x)_0 + u < f(x)_0 + \tau$.

This shows that we can always label the vertices on 01 with labels 0 or 1.

We may now apply Sperner's lemma : there is a full element of $\mathcal{T}_k$.

We will show that a vertex of the full triangle can only have a small displacement. Let the full triangle have vertices u, v, w. They have all distinct labels. In the picture below we have considered the labeling u — 0, v — 1, w — 2. The conditions that determined the labeling are $f(x)_i < x_i + \tau$; the dotted lines through u, v, w determine the part of Δ with points with 0-coordinate less than that of u, etc. The lines ℓ_u, ℓ_v, ℓ_w take into account the extra room given by the $+\tau$ (so they move away a bit from the full triangle). Since the vertices are less than δ apart, the images are less than ϵ apart. As the circle in the picture below shows, $f(u)$ cannot move up from ℓ_w more than ϵ, and neither move away to the right from ℓ_v more than ϵ. So $f(u)$ has to remain within the triangle bounded by ℓ_u, ℓ^v, and ℓ^w. One now immediately sees that the maximal distance between u and $f(u)$, as shown in the figure, is less than $2\delta + \tau + 2\epsilon$, and hence less than 4ϵ.

So for a given n, we can find point that moves less than 2^{-n+2}.



We have chosen in the proof a particular labeling. In total there are 6 cases of labelings to be considered. Our labeling is, so to speak, a worst case. In each case the analysis of the various admitted domains for the points $f(u), f(v), f(w)$ yields a distance between u and $f(u)$ that is not larger than in the above case.

The above method can be generalized to higher dimensional simplexes.

Given the similarity of the fixed point theorem and the intermediate value theorem, it is tempting to formulate a version of the fixed point theorem after the theorem in [Troelstra, A.S. and D. van Dalen 1988], p. 292 ff, Although it seems rather plausible, there is a yet no proof, so the best we can do is to put it as a conjecture : Let f be a continuous mapping of Δ into itself, with property that each open set contains a point x with $x \# f(x)$, then f has a fixed point. Looking at the one dimensional case, it seems advisable to add the condition that for x on the sides of Δ we have $x \# f(x)$.

Références

[Brouwer 1919] L.E.J. Brouwer. Intuitionistische Mengenlehre . *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 28 :203–208, 1919. Appeared in 1920.

[Brouwer 1952] L.E.J. Brouwer. An intuitionist correction of the fixed-point theorem on the sphere. *Proceedings of the Royal Society London*, 213 :1–2, 1952.

[Brouwer 1992] L.E.J. Brouwer. *Intuitionismus (*Ed. D. van Dalen). Bibliographisches Institut, Wissenschaftsverlag, Mannheim, 1992.

[Dalen, D. van 2008] Dalen, D. van. *Logic and Structure*. Springer, Berlin, 2008. 4th ed.

[Henle 1979] M.. Henle. *A Combinatorial Introduction to Topology.* W.H. Freeman, San Francisco, 1979.

[Soerensen, M.H. and P. Urzyczyn 2006] Soerensen, M.H. and P. Urzyczyn. *Lectures on the Curry-Howard Isomorphism*. Elsevier, Amsterdam, 2006.

[Troelstra, A.S. and D. van Dalen 1988] Troelstra, A.S. and D. van Dalen. *Constructivism in Mathematics, I, II.* North-Holland Publ. Co., Amsterdam, 1988.

Department of Philosophy
Utrecht University