Cantor-Bernstein implies Excluded Middle

Pierre Pradic [*]   Chad E. Brown[†]

April 22, 2019

Abstract

We prove in constructive logic that the statement of the Cantor-Bernstein theorem implies excluded middle. This establishes that the Cantor-Bernstein theorem can only be proven assuming the full power of classical logic. The key ingredient is a theorem of Martín Escardó stating that quantification over a particular subset of the Cantor space $2^{\mathbb{N}}$, the so-called one-point compactification of $\mathbb{N}$, preserves decidable predicates.

The *Cantor-Bernstein theorem* is an elementary statement $\mathbb{CB}$ of set theory: for any two sets A and B, if there are injections $f : A \hookrightarrow B$ and $g : B \hookrightarrow A$, then there exists a bijection $h : A \xrightarrow{\sim} B$. An interesting feature of this theorem is that it may be proven in ZF without assuming the axiom of choice. However, this proof is non-constructive in the sense that it goes through classical logic; while the construction of the bijection h is rather explicit, one needs to appeal to excluded middle to show that it is indeed a bijective function.

Models of constructive set theory invalidating $\mathbb{CB}$ are known, such as for instance, models based on Kleene realizability[1]. However, it left open the question of whether the full power of *excluded middle* ($\mathbb{EM}$) is really necessary to prove the theorem or if a weaker classical principle would be enough. The purpose of this note is to show that, indeed, full excluded middle is required because of the following theorem.

Theorem 1. ($\mathbb{CB} \Rightarrow \mathbb{EM}$) *Over intuitionistic set theory, the Cantor-Bernstein theorem implies the principle of excluded middle.*

The argument is a straightforward application of a key theorem of Martín Escardó [3] concerning the one-point compactification $\mathbb{N}_\infty$ of $\mathbb{N}$ and decidable predicates: there exists a function

$$\varepsilon \quad : \quad (\mathbb{N}_\infty \to 2) \quad \longrightarrow \quad \mathbb{N}_\infty$$

selecting a counter-witness for its input when possible. Formally speaking, it means that for any decidable predicate $P : \mathbb{N}_\infty \to 2$, we have the following equivalence[2].

$$\big(\forall p \in \mathbb{N}_\infty.\ P(p) = 1\big) \qquad \Longleftrightarrow \qquad P(\varepsilon(P)) = 1$$

This result is rather striking as it means that there exists an infinite set[3] for which decidable predicates are stable under quantification, *provably in constructive logic.* This is to be contrasted against the case of $\mathbb{N}$, which admits no recursive selection function.

Assuming the existence of *any* infinite set equipped with a selection function, $\mathbb{CB} \Rightarrow \mathbb{EM}$ may be proven by very elementary means. The reader familiar with [3] may content themselves with the proofs of Lemma 7 and Theorem 1. Section 3 as a whole gives a self-contained proof of $\mathbb{CB} \Rightarrow \mathbb{EM}$, integrating the necessary technical content from [3] about $\mathbb{N}_\infty$. For the more casual reader, we first give a preliminary example of an elementary set-theoretic statement implying excluded middle in Section 1 to show how the main theorem and many similar statements are proven by considering subsingleton sets. We then give an already-known proof [1] of a weaker variant of Theorem 1 in Section 2 for didactic purposes, before moving on to the proof of the main theorem.

[*]ENS de Lyon, Université de Lyon, LIP and University of Warsaw, MIMUW

[†]Czech Technical University in Prague

[1]For instance, consider the subset $H \subseteq \mathbb{N}$ corresponding to the halting problem. If $\mathbb{CB}$ held in the effective topos, we would have a recursive bijection $\mathbb{N} \xrightarrow{\sim} \{2n \mid n \notin H\} \cup \{2n + 1 \mid n \in \mathbb{N}\}$; since it is in particular surjective, we would be able to build a recursive enumeration of the complement of H, which is absurd.

[2]This should not be confused with Hilbert's ε which dually selects *witnesses* of existential statements.

[3]By which we systematically mean Dedekind-infinite here.

1

Foundational and notational preliminaries We work in a constructive set theory such as IZF or CZF where we do not assume excluded middle upfront. For our purpose, the distinction between these two systems will not be relevant. We shall reason informally and trust that the reader knowledgeable of other foundations be able to adapt the argument to other settings. We assume that the axioms of Zermelo's set theory hold. Among other things, this means that our universe of sets is closed under pairing, union, powerset (which we write $\mathcal{P}(-)$) and set comprehension. For sets A and B, function spaces B^A and disjoint unions $A + B$ are built as usual. We write $\mathsf{inl} : A \to A + B$ and $\mathsf{inr} : B \to A + B$ for the usual injections into disjoint unions. In particular, for every $x \in A + B$ there is either an $a \in A$ such that $x = \mathsf{inl}(a)$ or $b \in B$ such that $x = \mathsf{inr}(b)$. Let $0 = \emptyset$, $1 = \{0\}$ and $2 = \{0, 1\} \cong 1 + 1$.

In set theory, propositions or truth values can be arranged as a set $\Omega = \mathcal{P}(1)$, which is closed under all logical connective. For a formal proposition $p \in \Omega$, we sometimes abbreviate "$p = 1$" by "p" when writing formulas. *Excluded middle* $\mathbb{EM}$ may then be formally written as $\forall p \in \Omega.\ p \vee \neg p$[4]. Note that an equivalent formulation of $\mathbb{EM}$ in this setting is $\Omega \cong 2$, which is *not* the case in constructive settings. In the sequel, although they are definitionally the same in set theory, we often make a notational distinction between propositions $p \in \Omega$ and subsets $A \subseteq 1$ in order to keep the arguments more readable.

Finally, and crucially[5], we assume the axiom of infinity: there exists a Dedekind-infinite set $\mathbb{N}$ minimal for inclusion. We leave to the reader to disambiguate e.g. addition of numbers and disjoint union from context.

Related works We do not reprove $\mathbb{EM} \Rightarrow \mathbb{CB}$ here; while it is not necessary to read this note, they motivate a strengthening of $\mathbb{CB}$ metionned in Definition 4. Any introduction to set theory should have a satisfactory proof; for reference, one may look at Theorem 3.2 in [5]. For a historical perspective, the reader may be interested in [4]. The question of the nonconstructivity of $\mathbb{CB}$ from the categorical point of view was studied by Banaschewski and Brümmer in [1] and mentioned in Johnstone's Elephant [6] (Lemma D4.1.12, p. 950).

1   An elementary set-theoretic statement implying $\mathbb{EM}$

Let us start with an example of a lemma involving functions which may proved in an introduction to elementary set theory.

Proposition 2. *Let A and B be sets and $f : A \to B$ an injective function. Suppose that A is non-empty. If excluded middle holds, then there exists a surjection $g : B \to A$.*

Proof. Since A is non-empty, one can pick a default element $d \in A$. By excluded middle, we know for every $y \in B$ either there exists an $x \in A$ such that $f(x) = y$ or no such x exists. Hence we can define g by cases as follows:

$$g(y) := \begin{cases} x & \text{if } f(x) = y \\ d & \text{if no such } x \text{ exists} \end{cases}$$

Note g is well defined since f is injective, and g is clearly surjective. □

Notice that in this little proof, one needs to make a case analysis using excluded middle. As in the case of Cantor-Bernstein, one can ask if this is necessary. In fact, it is necessary as we now prove.

Proposition 3. *Suppose for all sets A and B there is a surjective function $g : B \to A$ whenever A is nonempty and there is an injective function $f : A \to B$. Then excluded middle holds.*

Proof. Let $p \in \Omega$ be given and consider $A := \{0 \mid p\}$. There is clearly an injection $f : A + 1 \to 2$ given by $f(\mathsf{inl}(x)) = 0$ for $x \in A$ and $f(\mathsf{inr}(0)) = 1$. Note that $A + 1$ is non-empty, as $\mathsf{inr}(0) \in A + 1$. Applying our assumption, we obtain a surjection $g : 2 \to A + 1$. Note that $\forall i \in 2.\ g(i) = \mathsf{inl}(0) \vee g(i) = \mathsf{inr}(0)$ holds because of the universal property of disjoint unions. Hence we have two cases.

---

[4]Note that the scheme $\forall x.\ \varphi(x) \vee \neg\varphi(x)$ for arbitrary φ is no more general since it can be recovered by set comprehension.

[5]Interestingly, one may easily construct models of finitary models satisfying $\mathbb{CB}$ but not $\mathbb{EM}$: take the internal logic of the topos $\mathsf{Finset}^{\mathcal{C}^{\mathrm{op}}}$. If $\mathcal{C}$ is not a groupoid, $\mathbb{EM}$ is not satisfied, while $\mathbb{CB}$ always hold in the internal logic. This indicates that assuming the axiom of infinity is essential here.

- Suppose $g(i) = \mathsf{inl}(0)$ for some $i \in 2$. In this case $0 \in A$ and so p holds.

- Suppose $g(0) = \mathsf{inr}(0)$ and $g(1) = \mathsf{inr}(0)$. We prove $\neg p$ holds. To this end, assume p holds. Hence $0 \in A$. Since g is surjective there must be some $i \in 2$ such that $g(i) = \mathsf{inl}(0)$, contradicting our assumption.

$\square$

Here, the strategy was fairly simple: take the $A \subseteq 1$ associated to the proposition, and try to make it fit in the hypothesis of the lemma using disjoint unions and singletons. The situation in the proof can be visualized as follows:
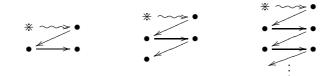

2 Reversing Cantor-Bernstein-Banaschewski-Brümmer

One can try to adopt a similar strategy for proving that Cantor-Bernstein implies excluded middle.

Provided some $A \subseteq 1$, one can start building an injection $f : A \to 1$.

However, we need to have an injection going back and we are unsure of the existence of an element in A. So let us consider the obvious injection $1 \to A + 1$.

Again we need a new value to be the image under f of this latest element, which leads us to consider 2. Since we still do not have two injections, one might be tempted to iterate this process.

This informal discussion suggests using $\mathbb{N}$ and the following injections.

$$f : \quad \mathbb{N} \longrightarrow A + \mathbb{N}$$
$$n \mapsto \mathsf{inr}(n)$$

$$g : \quad A + \mathbb{N} \longrightarrow \mathbb{N}$$
$$\mathsf{inl}(0) \mapsto 0$$
$$\mathsf{inr}(n) \mapsto n+1$$

$\mathbb{CB}$ then provides a bijection $h : \mathbb{N} \to A + \mathbb{N}$. In fact, in elementary proofs of the theorem this bijection can be seen as a perfect matching of the above graph. Note however that the usual statement $\mathbb{CB}$ conceals this relationship between f, g and h. Banaschewski and Brümmer [1] studied the corresponding strengthened version of $\mathbb{CB}$, which we dub $\mathbb{CBBB}$, in a categorical setting and proved that it implied excluded middle.

Definition 4. *We say $\mathbb{CBBB}$ holds if the following statement holds: given sets A and B and injections $f : A \to B$ and $g : B \to A$, there is a bijection $h : A \to B$ such that for all $x \in A$ and $y \in B$, $f(x) = y$ or $x = g(y)$ whenever $h(x) = y$.*

Let us remark that it is obvious that $\mathbb{CBBB} \Rightarrow \mathbb{CB}$. Let us also stress that $\mathbb{EM} \Rightarrow \mathbb{CBBB}$ can be easily obtained by adapting elementary proofs of $\mathbb{EM} \Rightarrow \mathbb{CB}$.

Theorem 5 (Proposition 4.1 in [1]). *If $\mathbb{CBBB}$ holds, then excluded middle holds.*

Proof. Assume $\mathbb{CBBB}$ holds. Let a proposition $p \in \Omega$ be given, seen as a subset $A = \{0 \mid p\} \subseteq 1$. Take $f : \mathbb{N} \to A + \mathbb{N}$ and $g : A + \mathbb{N} \to \mathbb{N}$ to be the injections described above.

By $\mathbb{CBBB}$ there is a bijection $h : \mathbb{N} \to A + \mathbb{N}$ such that $f(x) = y$ or $g(y) = x$ whenever $h(x) = y$ for $x \in \mathbb{N}$ and $y \in A + \mathbb{N}$. We know either $h(0) = \mathsf{inl}(0)$ or $h(0) = \mathsf{inr}(n)$ for some $n \in \mathbb{N}$.

- If $h(0) = \mathsf{inl}(0)$, then $0 \in A$ and so p holds.

- Suppose $h(0) = \mathsf{inr}(n)$ for some $n \in \mathbb{N}$. We will prove $\neg p$ holds. To this end, assume p holds, so that $0 \in A$. Since h is surjective, there is some $m \in \mathbb{N}$ such that $h(m) = \mathsf{inl}(0)$. Either $f(m) = \mathsf{inl}(0)$ or $g(\mathsf{inl}(0)) = m$. The first case is impossible since $f(m) = \mathsf{inr}(m)$ by the definition of f and $\mathsf{inl}(0) \neq \mathsf{inr}(m)$. Therefore, $g(\mathsf{inl}(0)) = m$ and, by definition of g, we must have $m = 0$. This is also impossible since it implies $\mathsf{inl}(0) = h(m) = h(0) = \mathsf{inr}(n)$.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

3   Reversing Cantor-Bernstein

Let us pause a moment and consider why we failed to prove the analogue of Proposition 3. In that proof of that proposition lemma, we did not use any information about the surjection g. Instead, we resorted to exhausitvely enumerating the set 2 to check whether we had some $x \in 2$ such that $g(x) = \mathsf{inr}(0)$, which is a decidable property. This feature of 2 of being searchable may be formalized using the notion of *omniscience*.

Definition 6 (Omniscient sets)**.** *We say a set O is* omniscient *if for every $p \in 2^O$ if either there exists $x \in O$ such that $p(x)$ is equal to 0, or p is constantly equal to 1. That is,*

$$\forall p \in 2^O.(\exists x \in O.p(x) = 0) \vee (\forall x \in O.p(x) = 1)$$

In classical logic, all sets are clearly omniscient, but this is not necessarily true in constructive logics. However, all finite sets, and in particular 2, are omniscient. This concept allows us to isolate the actual core of the proof of Proposition 3.

Lemma 7. *Suppose that we have an omniscient set O and some sets A and B. If there exists a surjection $f : O \to A + B$, then either A is inhabited or it is empty.*

Proof. Let $f : O \to A + B$ be given and define $p \in 2^O$ by

$$p(x) = \begin{cases} 0 & \text{if } \exists a \in A.\ f(x) = \mathsf{inl}(a) \\ 1 & \text{if } \exists b \in B.\ f(x) = \mathsf{inr}(b) \end{cases}$$

Since O is omniscient either $\exists x.\ P(x) = 0$ or $\forall x \in O.\ P(x) = 1$. If $\exists x \in O.\ P(x) = 0$, then A is clearly inhabited. Suppose $P(x) = 1$ for every $x \in O$. We will prove A is empty. Suppose $a \in A$. Since f is surjective there must be some $x \in O$ such that $f(x) = \mathsf{inl}(a)$, contradicting $P(x) = 1$. $\quad\square$

From $\mathbb{CB}$ (instead of the stronger $\mathbb{CBBB}$) we could use the injections from the proof of Theorem 5 to obtain a surjection $\mathbb{N} \to A + \mathbb{N}$. If $\mathbb{N}$ were omniscient, then we could use this surjection with Lemma 7 to $\mathbb{EM}$. However, omniscience of $\mathbb{N}$ correspond to the axiom of *limited principle of omniscience* ($\mathbb{LPO}$) a well-known constructive taboo [2], which can be thought of as a (strictly weaker) version of $\mathbb{EM}$[6]. This means that deriving $\mathbb{EM}$ from the existence of bijections $\mathbb{N} \cong 1 + \mathbb{N}$ by way of Lemma 7 is unreasonable. Luckily, Escardó proved that there exists an infinite subset of the Cantor space, $\mathbb{N}_\infty$, which is omniscient [3] and can be used to prove $\mathbb{CB} \Rightarrow \mathbb{EM}$.

In order to keep the argument self-contained, we reproduce his argument below before deriving the main result.

Definition 8. *We define $\mathbb{N}_\infty$ to be the set of non-increasing sequences in $2^{\mathbb{N}}$, i.e.[7],*

$$\mathbb{N}_\infty = \big\{ p \in 2^{\mathbb{N}} \bigm| \forall n \in \mathbb{N}.\ \big(p(n) = 1 \ \Rightarrow\ \forall m \in \mathbb{N}.\ (m < n \ \Rightarrow\ p(m) = 1)\big) \big\}.$$

---

[6] Remark that, at this point, we have $\mathbb{LPO} \wedge \mathbb{CB} \ \Rightarrow\ \mathbb{EM}$ over constructive set theory. This observation is however not necessary to carry out the subsequent argument.

[7]For more categorically-inclined people, $\mathbb{N}_\infty$ is the final coalgebra for the functor $X \mapsto 1 + X$. This justifies calling $\mathbb{N}_\infty$ the set of conatural numbers. The induced topology from $2^{\mathbb{N}}$ in Definition 8 also justifies calling $\mathbb{N}_\infty$ the *one-point compactification of* $\mathbb{N}$.

Let $\omega \in \mathbb{N}_\infty$ denote the constant 1 function. We define an injection taking $n \in \mathbb{N}$ to $\underline{n} \in \mathbb{N}_\infty$ by taking

$$\underline{n}(m) = \begin{cases} 1 & \text{if } m < n \\ 0 & \text{otherwise.} \end{cases}$$

Finally we define $\underline{S} : \mathbb{N}_\infty \to \mathbb{N}_\infty$ by cases taking $\underline{S}(p)(0) = 1$ and $\underline{S}(p)(n+1) = p(n)$.

The set $\mathbb{N}_\infty$ is infinite as witnessed by $\underline{0}$ and $\underline{S}$.

Lemma 9. *The function $\underline{S} : \mathbb{N}_\infty \to \mathbb{N}_\infty$ is injective and $\underline{S}(p) \neq \underline{0}$ for all $p \in \mathbb{N}_\infty$.*

Proof. Suppose $\underline{S}(p) = \underline{S}(q)$. Since $p(n) = \underline{S}(p)(n+1) = \underline{S}(q)(n+1) = q(n)$ for every $n \in \mathbb{N}$ we know $p = q$, as desired. The fact that $\underline{S}(p) \neq \underline{0}$ for all $x \in \mathbb{N}_\infty$ follows from $\underline{S}(p)(0) = 1 \neq 0 = \underline{0}(0)$. $\square$

Classically every element of $\mathbb{N}_\infty$ is either ω or of the form $\underline{n}$. The corresponding disjunction is equivalent to $\mathbb{LPO}$, and so is unprovable constructively. However, for decidable predicates, it is sufficient to show that they hold over all elements $\underline{n}$ and ω to show they hold everywhere[8].

Lemma 10. *Let $Q \in 2^{\mathbb{N}_\infty}$ be given. If $Q(\omega) = 1$ and $\forall n \in \mathbb{N}. \, Q(\underline{n}) = 1$, then $\forall p \in 2^{\mathbb{N}_\infty}. \, Q(p) = 1$.*

Proof. Let $Q \in 2^{\mathbb{N}_\infty}$ such that $Q(\omega) = 1$ and $\forall n \in \mathbb{N}. \, Q(\underline{n}) = 1$. Let $p \in \mathbb{N}_\infty$ be given. To prove $Q(p) = 1$, it is enough to prove $Q(p) \neq 0$. Assume $Q(p) = 0$. Under this assumption we can prove $\forall n \in \mathbb{N}. \, p(n) = 1$ by strong induction. Assume $\forall k \in \mathbb{N}. \, (k < n \Rightarrow p(k) = 1)$ and $p(n) = 0$. This is enough information to infer $p = \underline{n}$, contradicting $Q(p) = 0$ and $Q(\underline{n}) = 1$. To end the proof we note that p must be ω (since $\forall n \in \mathbb{N}. \, p(n) = 1$), contradicting $Q(p) = 0$ and $Q(\omega) = 1$. $\square$

While the desired function $\varepsilon : 2^{\mathbb{N}_\infty} \to \mathbb{N}_\infty$ is rather straightforward to define, Lemma 10 is critical in allowing to prove constructively that it is indeed a selection function.

Theorem 11 (Theorem 3.15 in [3]). *There is a function $\varepsilon : 2^{\mathbb{N}_\infty} \to \mathbb{N}_\infty$ such that for every $Q \in 2^{\mathbb{N}_\infty}$, if $Q(\varepsilon(Q)) = 1$, then $\forall p \in 2^{\mathbb{N}_\infty}. \, Q(p) = 1$.*

Proof. For $Q \in 2^{\mathbb{N}_\infty}$, take $\varepsilon(Q) \in 2^{\mathbb{N}}$ to be

$$\varepsilon(Q)(n) = \begin{cases} 1 & \text{if } Q(\underline{k}) \text{ for each } k \leq n \\ 0 & \text{otherwise.} \end{cases}$$

This may be well-defined recursion over n. It is easy to check that $\varepsilon(Q) \in \mathbb{N}_\infty$ as well.

Assume $Q(\varepsilon(Q)) = 1$ and let $p \in \mathbb{N}_\infty$ be given. If $\forall k < n. \, Q(\underline{k}) = 1$ and $Q(\underline{n}) = 0$, then $\varepsilon(Q) = \underline{n}$ and so $Q(\underline{n}) = Q(\varepsilon(Q)) = 1$, contradicting $Q(\underline{n}) = 0$. Consequently, $Q(\underline{n}) = 1$ for every n by induction. Thus $\varepsilon(Q) = \omega$ and so $Q(\omega) = Q(\varepsilon(Q)) = 1$ holds as well. Hence $Q(p) = 1$ for every $p \in 2^{\mathbb{N}}$ by Lemma 10. $\square$

Corollary 12 (Corollary 3.6 in [3]). *The set $\mathbb{N}_\infty$ is omniscient.*

Proof. Let $Q \in 2^{\mathbb{N}_\infty}$ be given. If $Q(\varepsilon(Q)) = 0$, then $\exists x \in \mathbb{N}_\infty. Q(x) = 0$. If $Q(\varepsilon(Q)) = 1$, then $\forall x \in \mathbb{N}_\infty. \, Q(x) = 1$ by Theorem 11. $\square$

This completes the part of the construction we obtained following Escardó [3]. We can now easily put it together with Lemma 7 to conclude.

Proof of Theorem 1. Assume $\mathbb{CB}$ holds. We know $\mathbb{N}_\infty$ is omniscient by Lemma 12. We know $\underline{S}$ is injective and $\underline{S}(x) \neq \underline{0}$ for every $x \in \mathbb{N}_\infty$ by Lemma 9. Let $p \in \Omega$ be a proposition and take $A = \{0 \mid p\} \subseteq 1$. Analogously to Section 2, we consider the following functions:

$$
\begin{array}{llll}
f: & \mathbb{N}_\infty & \longrightarrow & A + \mathbb{N}_\infty \\
 & x & \mapsto & \mathsf{inr}(x)
\end{array}
\qquad
\begin{array}{llll}
g: & A + \mathbb{N}_\infty & \longrightarrow & \mathbb{N}_\infty \\
 & \mathsf{inl}(0) & \mapsto & \underline{0} \\
 & \mathsf{inr}(x) & \mapsto & \underline{S}(x)
\end{array}
$$

Both f and g are clearly injective, so we can apply $\mathbb{CB}$ to obtain a bijection $h : \mathbb{N}_\infty \to A + \mathbb{N}_\infty$. Lemma 7 now implies that either A is inhabited (so p holds) or A is empty (so $\neg p$ holds). $\square$

---

[8]This constitutes a particular case of Lemma 3.4 in [3].

References

[1] G. C.L. Banaschewski, B.; Brümmer. Thoughts on the Cantor-Bernstein Theorem. *Quaestiones Mathematicae*, 9, 01 1986.

[2] E. Bishop. *Foundations of constructive analysis*. McGraw-Hill, 1967.

[3] M. Escardó. Infinite sets that satisfy the principle of omniscience in any variety of constructive mathematics. *J. Symb. Log.*, 78(3):764–784, 2013.

[4] A. Hinkis. *Proofs of the Cantor-Bernstein Theorem: A Mathematical Excursion*. Science Networks. Historical Studies, Vol. 45. Birkhäuser (Springer Basel), 2013.

[5] T. Jech. *Set theory*. Springer Science & Business Media, 2013.

[6] P. T. Johnstone. *Sketches of an Elephant: A Topos Theory Compendium: 2 Volume Set*. Oxford University Press UK, 2002.