Double Closure Induction and a Theorem of Hessenberg and Bourbaki

Walter Felscher

1962

This is a translation of the paper *Doppelte Hülleninduktion und ein Satz von Hessenberg und Bourbaki* by Walter Felscher that was published in *Archiv der Mathematik* **13** (1962) 160–5.

The paper gives the history of the so-called Bourbaki–Witt theorem, tracing it back to Gerhard Hessenberg (1908) and Kazimierz Kuratowski (1922). However, all of the versions of this proof seem to be terribly laboured, so the value of this paper lies in the history rather than the proof itself.

The translation was made with the help of ImageMagick, Tesseract OCR, the DeepL translator and a lot of hand-editting.

Felscher posted the following to the *Historia Mathematica* email forum on 17 April 2000, but sadly died later that year:

`archives.math.utk.edu/hypermail/historia/apr00/0105.html`

"Returning to Hessenberg, his paper

Kettentheorie und Wohlordnung. Crelle 135 (1909) 81-133

can hardly be underestimated in its importance. Not that it was understood by his contemporaries. But Hessenberg, analyzing Zermelo's second proof of the well-ordering theorem, studied the general ways to construct well ordered subsets of ordered sets — with the one restriction that order always was inclusion and ordered sets were subfamilies of power sets. In the course of this, Hessenberg stated and proved the fixpoint theorem which thirty years later was rediscovered — for ordered sets now — by Nicolas Bourbaki. The amazing thing is that Hessenberg's proof is precisely the same as that given by Bourbaki! (only that at one small point a simpler argument can be used due to the circumstance that Hessenberg's order is inclusion). For details, I refer to my article in Archiv d.Math. 13 (1962) 160-165 and to my book Naive Mengen und Abstrakte Zahlen from 1979, p.200 ff."

Theorem 0.1 Let $(E, \leq)$ be a poset (partially ordered set) and $f : E \to E$ be an order-preserving function such that

- E has a least element e;
- E has suprema of al chains; and
- $\forall x \in E.\, x \leq fx$.

(Recall that a chain $C \subset E$ is a subposet in which $\forall xy \in C.\, x \leq y \vee y \leq x$.)

Let $D \subset E$ be the smallest subset such that

- $e \in D$;
- $\forall x \in E.\, x \in D \Rightarrow fx \in D$; and
- D is closed under suprema of chains.

Then D has the *key property* that

$$\forall x, y \in D. \quad x \leq y \ \vee \ fy \leq x,$$

whence it is itself a chain. Hence D has a supremum, which is the least fixed point of f.

Moreover D is well founded, in the sense that any non-empty subset $U \subset D$ has a $\leq$-least element.

Dedicated by the author to Reinhold Baer on his 70th birthday.
Received 10 November 1961

Mathematisches Institut, Universität Freiburg, Hebelstraße 40

Let $(E, \leq)$ be a partially ordered set, f a mapping of E into itself with $x \leq fx$ for all $x \in E$ and $d \in E$. Then let D be the smallest subset of E that contains d, is closed under f and is closed under any suprema that exist of nonempty subsets of E.

Then $(D, \leq)$ is totally and indeed well-ordered.

In the following note we give a proof of this well known theorem that is not really new: a proof only, which the known methods of Dedekind's chain theory using the evidence of a general principle of to a general principle of double closure induction.

⟦Dedekind: Was sind und was sollen die Zahlen?⟧

A first special case of this theorem can already be found in the idea of Zermelo's second proof of the well-ordering theorem. The ordered set $(E, \leq)$ has the form $(\mathcal{P}M, \subset)$, suprema are unions of sets and for f it follows from $x \neq fx$ that fx is the upper neighbour of x. The set fx contains exactly one element more than the set x (Zermelo [14]; there, however, a dual formulation is used, with $\supset$ instead of $\subset$).

In the same year 1908 that Zermelo's proof appeared, the theorem was also proved by Hessenberg (Hessenberg [4], pp. 127 to 129), again for an ordered set $(\mathcal{P}M, \subset)$, but now for any mapping f from $\mathcal{P}M$ to itself with $X \subset fX$ for all $X \in \mathcal{P}M$. Hessenberg proof was reproduced, somewhat modified, by Kuratowski (Kuratowski [7], Th. 3).

Formulated and proved in the form stated at the beginning, for arbitrary ordered sets, the theorem is then found in the almost simultaneously published works of by Bourbaki (Bourbaki [2]), Witt (Witt [12]) and Inagaki (Inagaki [5]). Not only do their proofs agree with his (apart from an additional but dispensable condition in Inagaki's), but they also resemble Hessenberg's, step by step, except that now we speak of $\leq$ instead of $\subset$ and of suprema instead of unions of sets. The abstraction from sets of the form $(\mathcal{P}M, \subset)$ to general ordered sets does not require any new proof ideas.

And this theorem of Bourbaki, Witt and Inagaki then became widely known, namely as an auxiliary result in the derivation of Zorn's Lemma from the Axiom of Choice.

Even before Bourbaki, however, Milgram had already proved a theorem in 1939 (Milgram [8]) from which the validity of the one here follows even under still more general conditions: it is sufficient if D, instead of being closed under suprema, is closed against a function g that assigns upper bounds to certain subsets of E in such a way that subsets which generate equal beginnings also have equal image and further the g-image of a set with maximum is equal to this maximum or equal to the f-image of the maximum.

Milgram refines Hessenberg's approach in his proof; it would be shown that Milgram's methods cannot be easily captured by the induction methods considered here.

Later Vaughan (Vaughan [11]) and Banaschewski (Banaschewski [1]), linked to H. Kneser (Kneser [6]) and Szele (Szele [10]), it was shown that results of this kind can also be derived according to the ideas of Zermelo's first proof of the well-ordering theorem (Zermelo [13]).

A comprehensive analysis of how well-ordered sets can be constructed from given mappings will be published elsewhere (Felscher [13]).

In this paper, we first establish a principle of double closure induction. from which a somewhat simplified form of the Hessenberg–Bourbaki proof of that theorem can be immediately obtained as a special case. That something like this was to be expected follows from a remark by J. Schmidt (Schmidt [9], p. 144), who recognised from a certain principle of double closure induction that it played a role in parts of the Zermelo–Bourbaki proofs. The double closure induction considered here is nothing else than an elaboration of the double closure induction of J. Schmidt.

In a first supplementary remark it is then shown, that also the double closure induction is not, as it will be done first, proved *ad hoc*, but by similar systematic considerations as the double closure induction. In two further remarks, the special cases and generalizations of the theorem.

1

Let E be a set.

Suppose we have a partial mapping $g : \mathcal{P}E \rightharpoonup E$, where we write $\mathsf{def}\, g$ for the domain of definition of g. We say that a subset $M \subset E$ is ***g-closed*** if whenever $B \subset M$ with $B \in \mathsf{def}\, g$ we have $gB \in M$.

⟦We think of g as supremum, but he's trying to avoiding that specificity. He should nevertheless use a more suggestive symbol than a letter.⟧

Let R be a subset of the pair-set $E \times E$, although instead of $(x, y) \in R$ we write xRy. For $x \in E$ let Rx be the set of all $y \in E$ with xRy and $R^{-1}x$ the set of all $y \in E$ with yRx.

⟦It would be more logical to write $xR \equiv \{y \mid xRy\}$ instead of Rx and $Ry \equiv \{x \mid xRy\}$ instead of $R^{-1}y$.⟧

Let A be a subset of E such that Ra is g-complete for all $a \in A$.

Then the set V of all $x \in E$ with $A \subset R^{-1}x$ is also g-complete.

For from $B \subset V$, so $A \subset R^{-1}b$ for all $b \in B$, it follows that aRb for all $b \in B$ and $a \in A$, therefore $b \in Ra$ for all $a \in A$ and all $b \in B$.

From $B \in \mathsf{def}\, g$ and the g-completeness of each Ra it follows that $gB \in Ra$ for all $a \in A$, hence $a \in R^{-1}gB$ for all $a \in A$, hence $A \subset R^{-1}gB$, so $gB \in V$. ⟦correct⟧

Now let $f : E \to E$. We say that a subset $M \subset E$ is ***f*-closed** if $\forall x.\, x \in M \Rightarrow fx \in M$.

We write $\mathfrak{H}_g$ and $\mathfrak{H}_f$ for the systems of all g-closed and f-closed subsets of E, respectively.

These systems of subsets are closed under intersection and in particular they contain E. Moreover, $\mathfrak{H}_g \cap \mathfrak{H}_f$ is another such a system, consisting of the subsets of E that are both g- and f-closed.

Suppose that $\emptyset$ is not in $\mathsf{def}\, g$ ⟦why?⟧, let $d \in E$, and let D be the **(g, f)-hull** of $\{d\}$, that is, the intersection of all sets from $\mathfrak{H}_g \cap \mathfrak{H}_f$ that contain d. Again let R be a subset of $D \times D$.

The definition of D leads immediately to the following simple principle of induction:

R is equal to $D \times D$ if the set V of all $x \in D$ with $R^{-1}x = D$ contains the element d, is g-closed and f-closed.

We claim that following double induction principle is valid:

R is equal to $D \times D$ so long as

(i) $R^{-1}d = D$ ⟦$\forall x \in D.\, xRd$⟧,

(ii) *for all $x, y \in D$, if xRy and yRx, then $xRfy$,*

(iii) *for all $x \in D$, $Rx \in \mathfrak{H}_g$ ⟦if $B \in \mathsf{def}\, g$ and $\forall y \in B.\, xRy$ then $xR(gB)$⟧.*

First, for all $x \in D$: if $R^{-1}x = D$, then $Rx = D$ ⟦if $\forall y \in D.\, yRx$ then $\forall y \in D.\, xRy$⟧, because

(i) gives $d \in Rx$ ⟦xRd⟧;

(ii) if $y \in Rx$, that is, xRy, then since $R^{-1}x = D$ ⟦$\forall y \in D.\, yRx$⟧ also yRx, whence $xRfy$ ⟦by hypothesis (ii)⟧ *i.e.* $fy \in Rx$, giving f-closure.

(iii) gives g-closure ⟦if $B \in \mathsf{def}\, g$ and $\forall y \in B.\, xRy$ then $xR(gB)$⟧.

And now we show that the set V of all $x \in D$ with $R^{-1}x = D$ ⟦$V \equiv \{x \in D \mid \forall y \in D.\, yRx\}$, my S⟧ is equal to D:

(i) gives $d \in V$ ⟦$\forall y \in D.\, yRd$⟧;

(ii) from $R^{-1}x = D$ follows $Rx = D$ ⟦if $x \in V$, *i.e.* $\forall y \in D.\, yRx$ then $\forall y \in D.\, xRy$⟧, thus for all $y \in D$ and xRy; if one exchanges x and y in (ii), then $yRfx$ for all $y \in D$, $R^{-1}x = D$ ⟦$\forall y \in D.\, \forall y \in D.\, yRx$ eh??⟧, giving f-closure ⟦$\forall y \in D.\, yR(fx)$ *i.e.* $fx \in V$⟧.

(iii) gives g-closure ⟦if $B \in \mathsf{def}\, g$ and $\forall y \in B.\, \forall x \in D.\, xRy$ then $\forall x \in D.\, xR(gB)$, *i.e.* $gB \in V$⟧. □

⟦I cannot fault this proof, but the notation and choice of lettering are confused and I remain suspicious of the case $(gB)R(fx)$.⟧

In the case $g = \emptyset$ one obtains from this the principle of double complete induction of J. Schmidt [1] and the proof of it is found there, p. 145.

Let $(E, \leq)$ be an ordered set and f a mapping from E into itself with $x \leq fx$ for all $x \in E$. Let g be a mapping from $\mathcal{P}E$ into E such that if $B \in \mathsf{def}\, g$ and B is not is empty, then it has a supremum that is just gB. Further let $d \in E$ and D be the (g, f)-hull of $\{d\}$.

Then $(D, \leq)$ is totally ordered.

Consider the relation R defined by

$$xRy \quad \textit{exactly if} \quad x, y \in D \textit{ and } (y \leq x) \vee (fx \leq y).$$

Since $x \leq fx$, it follows from $fx \leq y$ that also $x \leq y$. It therefore suffices to prove $R = D \times D$.

(i) holds, because $d \leq x$ by construction of D for all $x \in D$;

(ii) holds: suppose xRy, so $y \leq x$ or $fx \leq y$ ⟦we also assume $yRx \equiv x \leq y \vee fy \leq x$⟧;

- if $y \leq x$ but $y \neq x$ then $x \not\leq y$, so from yRx we have $fy \leq x$ and $xRfy$;
- if $y = x$ then $fy = fx$, so $fx \leq fy$ and $xRfy$;
- if $fx \leq y$, since $y \leq fy$ we have $fx \leq fy$ and $xRfy$.

⟦If f preserves order then $xRy \Rightarrow yR(fx)$ very simply.⟧

(iii) from $x \in D$, $B \in \mathsf{def}\, g$, $B \subset Rx$ it follows that xRb, so $(b \leq x) \vee (fx \leq b)$ for all $b \in B$;

- if $b \leq x$ for all $b \in B$, then also $gB = \sup B \leq x$;
- if $fx \leq b$ for some $b \in B$, then also $fx \leq b \leq \sup B = gB$.

Hence we always have $xRgB$ and $gB \in Rx$.

For $g = \emptyset$ ⟦*i.e.* without the infinitary operation g whose only meaningful example is join $\bigvee$⟧ this proof is found again in J. Schmidt [1], p. 146.

That $(D, \leq)$ is in fact well ordered, one now concludes from $R = D \times D$ in the usual way (cf. Hessenberg [1], Bourbaki [1]).

2

The scheme of the double closure induction formulated in section 1 can also be justified in the following way, where E, f, g, d, D, R and V are as explained there.

$V = D$ is proved if

(1) $R^{-1}d = D$,

(2) for all $x \in D$, if $R^{-1}x = D$, then $R^{-1}fx = D$;

(3) for all $x \in D$, $Rx \in \mathfrak{H}_g$.

(2) is proved if

(2a1) for all $x \in D$, if $R^{-1}x = D$ then $dRfx$,

(2a2) for all $x \in D$, if $R^{-1}x = D$ and $yRfx$ then $fyRfx$,

(2a3) for all $x \in D$, if $R^{-1}x = D$ then $R^{-1}fx \in \mathfrak{H}_g$,

and for this it is sufficient if
(2b1) $Rd = D$,
(2b2) for all $x, y \in D$, if $fyRx$ and $yRfx$ then $fxRfy$, (2a3).

Since $R = D \times D$ is equivalent to $R \cap R^{-1} = D \times D$, because of $(R \cap R^{-1})x = Rx \cap R^{-1}x$ the induction can also be carried out over $Rx \cap R^{-1}x = D$ instead of over $R^{-1}x = D$.

(For $g = \emptyset$ this can be found completely corresponding to the following J. Schmidt [1]), but then instead of (3) one would have the only slightly handy condition $Rx \cap R^{-1}x \in \mathfrak{H}_g$, for all $x \in D$.

But since from $Rfx \cap R^{-1}fx = D$ it also follows that $R^{-1}fx = D$, one can pass in (2) from R to $R \cap R^{-1}$ and thus have the induction scheme
(1) ?
(3) ?
(4) for all $x \in D$, if $R^{-1}x = D$ then $Rx \cap R^{-1}x = D$,
(5) for all $x \in D$, if $Rx \cap R^{-1}x = D$ then $Rfx \cap R^{-1}fx = D$.
According to what was said for (2), it is now sufficient for (5) if
(5b1) $Rd \cap R^{-1}d = D$,
(5b2) for all $x, y \in D$, if $fyRx$, $xRfy$, $yRfx$ and $fxRy$ then $fxRfy$ and $fyRfx$,
(5b3) for all $x \in D$, if $Rx \cap R^{-1}x = D$ then $Rfx \cap R^{-1}fx \in \mathfrak{H}_g$. —

Since the preconditions of (5b2) merge when x and y are interchanged, one can omit the second condition; if one continues to use only those conditions in which fx occurs, and one demands the validity of the condition thus arising with x in place of fx, then from (5b2) one finally obtains
(6) *for all $x, y \in D$, if yRx and xRx then $xRfy$,*
so just (ii) from section 1.

Thus (5b1) proves to be dispensable, because due to (1) $d \in Rd$, and $Rd \in \mathfrak{H}_g$ because of (3), and $Rd \in \mathfrak{H}_g$ because of (1) and (6).

Further (5b3) becomes dispensable: because of (3) it suffices there to show $R^{-1}fx \in \mathfrak{H}_g$, but with (6) $Rx \cap R^{-1}x = D$ then $R^{-1}fx = D$ even entails $R^{-1}fx = D$.

Finally, in (4) one needs only $Rx = D$, so that (4) is proved if
(41) for all $x \in D$, if $R^{-1}x = D$ then xRd,
(42) for all $x \in D$, if $R^{-1}x = D$ and xRy then $xRfy$,
(43) for all $x \in D$, if $R^{-1}x = D$ then $Rx \in \mathfrak{H}_g$;
and here follows: (41) from (1), (42) from (6), (43) from (3): (4) is also superfluous. The henceforth consisting of (1), (3), (6) is just the induction scheme that we considered in section 1.

3

Let $(E, \leq)$ be an ordered set, f, g, D and R be as in section 1. The scheme given there by induction over the two places of R agrees nearly with the proofs of Hessenberg, Bourbaki and Witt.

The only difference is that besides R still Q with

$$xQy \quad \textit{exactly if} \quad x, y \in D \textit{ and } (x \leq y) \wedge (x \neq y) \Rightarrow (fx \leq y)$$

is considered, where from $R^{-1}x = D$ it follows at once that $Q^{-1}x = D$, then first from $Q^{-1}x = D$ to $Rx = D$, and thus $Q^{-1}x = D$ is proved for all $x \in D$.

Inagak considers C with

$$xCy \quad \textit{exactly if} \quad x, y \in D \textit{ and } (x \leq y) \vee (x \leq y),$$

where $R \subset C$, concludes from $Q^{-1}x \cap C^{-1}x = D$ that $Rx = D$ and thus proves $Q^{-1}x \cap C^{-1}x = D$ for all $x \in D$.

Kuratowski's version of Hessenberg's proof uses P with

$$xPy \quad \textit{exactly if} \quad x, y \in D \textit{ and } (y \leq x) \Rightarrow (Ry = D),$$

and he proves by induction that $Px = D$ for all $x \in D$: where $Pfx = D$ from $Px = D$ is deduced by induction from $Px = D$ to $Rfx = D$. If one has $fxRy$ then in the case $ffx \leq y$ certainly $fxRfy$, in the case $y \leq fx$ and $y \neq fx$ because of $Px = D$, $Rx = D$, xRy also $y \leq x$, again because of $Px = D$, $Ry = D$ also $yRfx$, therefore according to (ii) $fxRfy$.

Zermelo, on the other hand, can operate with the stronger presupposition that fx is the upper neighbour of x; but then $C^{-1}x = D$ has so $Q^{-1}x = D$, so that from $C^{-1}x = D$ one can conclude that $Rx = D$, and thus thus prove $C^{-1}x = D$ for all $x \in D$.

4

From a theorem of Milgram (Milgram [1], Th. 1) it is easy to derive the following statement:

Let $(E, \leq)$ be an ordered set, f a mapping from E into itself with $x \leq fx$ for all $x \in E$, g a mapping from $\mathcal{P}E$ into E such that for $B \in \mathsf{def}\, g$ then gB is an upper bound of B, further from $B_1, B_1 \in \mathsf{def}\, g$ and B_1 is co-initial with B_2 also follows $gB_1 = gB_2$.

(Here B_1 and B_2 are called ***co-initial*** if every element of B_i lies under some element from B_j for $i, j = 1, 2$).

Finally $B \in \mathsf{def}\, g$, $\mathsf{max}\, B \neq \emptyset$ always implies $gB \in \mathsf{max}\, B \cup f(\mathsf{max}\, B)$.

Further, let $\emptyset \in \mathsf{def}\, g$, $d = g\emptyset$, and D be the (g, f)-hull of $\emptyset$. *Then* $(D, \leq)$ *is well-ordered.*

First of all, we see that the proof given in Section 1 cannot be easily applied to this general case: that Rx is g-closed can in any case no longer be justified in the way described there.

But if one defines

$$xSy \quad \textit{exactly if} \quad x, y \in D \textit{ and for all } B, \textit{ if } B \in \mathsf{def}\, g,\ x = gB, \textit{ so } (x \leq y) \vee (\exists b \in B.\, y \leq b),$$

then an induction proof can be given in the following steps:

(a) $R^{-1}d = D$,

(b) $S^{-1}d = D$,

(c) for all $x \in D$, if $S^{-1}x = D$ then $Rx \in \mathfrak{H}_g$,

(d) for all $x \in D$, if $R^{-1}x = D$ then $Rx \in \mathfrak{H}_f$,

(e) for all $x \in D$, if $Rx = D$ and $S^{-1}x = D$ then $(S^{-1} \cap R^{-1})fx = D$,

(f) for all $M \in \mathsf{def}\, g$, if for all $m \in M$ we have

$$(S^{-1} \cap R^{-1})m = D \quad \text{then} \quad SgM = D,$$

for all $M \in \mathsf{def}\, g$, if $SgM = D$, then $(S^{-1} \cap R^{-1})gM = D$.

Again, one uses more conveniently in the premises the statement following from $R^{-1}x = D$ statement $Q^{-1}x = D$ and the statement $S^{-1}x = D$ following from $T^{-1}x = D$ with

xTy exactly if $x, y \in D$ and for all B, if $B \in \mathsf{def}\, g$, $x = gB$ and for all $b \in B$, $b \leq y$, $b \neq y$ then $x \leq y$.

The difficulty of the proof lies in step (f): one first proves by induction that special case of $SgM = D$ which arises when specializing B to M in S; the results of the earlier induction steps are used. For details and classification in more comprehensive methods, we refer to the forthcoming work forthcoming work of Felscher [3].

Bibliography

[1] B. Banaschewski, Über die Konstruktion wohlgeordneter Mengen. Math. Nachr.10, 239–245 (1953).

[2] N. Bourbaki, Sur le théorème de Zorn. Arch. Math.2, 434–437 (1949), DOI 10.1007/bf02036949, sci-hub.st/10.1007/BF02036949.

[3] W.Felscher, Konstruktion wohlgeordneter Mengen und Dedekindsche Kettentheorie. Erscheint in Z. math. Logik Grundl. Math.

[4] G. Hessenberg, Kettentheorie und Wohlordnung. J. reine angew. Math.135, 81–133 (1909). ⟦download/papers/Hessenberg/chain-theory.pdf⟧

[5] T. Inagaki, Sur deux théorèmes concernant un ensemble partiellement ordonné. Math. J. Okayama Univ.1, 167–176 (1952). www.math.okayama-u.ac.jp/mjou/mjou1-46/mjou_pdf/ ⟦download/papers/Inagaki/Bourbak-Witt.pdf⟧

[6] H. Kneser, Eine direkte Ableitung des Zornschen Lemmas aus dem Auswahlaxiom. Math. Z.53, 110–113 (1950).

[7] K. Kuratowski, Une méthode d'élimination des nombres transfinis des raisonnements mathématiques. Fundamenta Math.3, 76–108 (1922). DOI 10.4064/fm-3-1-76-108. ⟦dowload/papers/Kuratowski/eliminate-ordinals.pdf⟧

[8] A. N. Milgram, Partially Ordered Sets, Separating Systems and Inductiveness. Rep. Math. Colloquium, II. Ser., issue1, 18–30 (1939). brief review at www.ams.org/journals/bull/1939-45-11/S0002-9904-1939-07069-9/S0002-9904-1939-07069-9.pdf

[9] J. Schmidt, Einige Prinzipien der doppelten Induktion. Math.-phys. Semesterber.6, 137 bis 147 (1958).

[10] T. Szele, On Zorn's Lemma. Publ. Math., Debrecen1, 254–6 (1950).

[11] H. E. Vaughan, Well-ordered subsets and maximal members of ordered sets. Pacific J. Math.2, 407–412 (1952). projecteuclid.org/journals/pacific-journal-of-mathematics/volume-2/issue-3/Well-ordered-subsets-and-maximal-members-of-ordered-sets/pjm/1103051786.full ⟦download/papers/Vaughan/well-ordered-subsets.pdf⟧

[12] E. Witt, Beweisstudien zum Satz von M. Zorn. Math. Nachr.4, 434–8 (1951), DOI 10.1002/mana.3210040138, onlinelibrary.wiley.com/doi/abs/10.1002/mana.3210040138

[13] E. Zermelo, Beweis, da jede Menge wohlgeordnet werden kann. Math. Ann.59, 514–6 (1904).

[14] E. Zermelo, Neuer Beweis für die Mglichkeit einer Wohlordnung. Math. Ann.65, 107–128 (1908), English translation in van Heijenoort, pages 183–198.