By contracting quantifiers and possibly adding “dummy” variables and expressions like $x_i = x_i$, we can assume that a given formula $\varphi$ is of the form

\[(0.1)\quad \exists x_1 \forall x_2 \ldots Q x_r \psi(\bar{y}, x_1, \ldots, x_r)\]

or

\[(0.2)\quad \forall x_1 \exists x_2 \ldots Q x_r \psi(\bar{y}, x_1, \ldots, x_r),\]

where $Q$ is either $\exists$ or $\forall$, and $\psi$ is quantifier-free. In the following, we focus on the form given in (0.1). The argument for the other form is similar.

With any $\varphi(\bar{y})$ in prenex normal form (0.1) we associate a $\Delta_0$ formula $\varphi^*(\bar{y}, z_1, \ldots, z_r)$ given as

$$\exists x_1 < z_1 \forall x_2 < z_2 \ldots \exists x_r < z_r \psi(\bar{y}, x_1, \ldots, x_r, z_1, z_2, \ldots, z_r).$$

Claim: For any formula $\varphi$ in prenex normal form, for any $\bar{a} \in N$, and any $i_0 < i_1 < i_2 < \ldots < i_r$ with $\bar{a} < b_{i_0}$,

\[(0.3)\quad N \models \varphi[\bar{a}] \iff M \models \varphi[\bar{a}, b_{i_1}, \ldots, b_{i_r}].\]

The claim is proved by induction on the formula length (see also Lemma 4.47, where this technique was first described). If $\varphi$ has no quantifiers at all, the claim is clear. So assume now $\varphi(\bar{y})$ is as in (0.1) with $r \geq 1$. Then the claim is that $\varphi[\bar{a}]$ holds in $N$ if and only if

$$\exists x_1 < b_{i_1} \forall x_2 < b_{i_2} \ldots Q x_r < b_{i_r} \psi(\bar{a}, x_1, \ldots, x_r)$$

holds in $M$.\(^1\)

The formula $\varphi^*(\bar{y}, z_1, \ldots, z_r)$ is

$$\exists x_1 < z_1 \forall x_2 < z_2 \ldots \exists x_r < z_r \psi(\bar{y}, x_1, \ldots, x_r, z_1, z_2, \ldots, z_r).$$

Let $\theta(\bar{y}, x_1)$ be

$$\forall x_2 \ldots Q x_r \psi(\bar{y}, x_1, \ldots, x_r),$$

so $\varphi(\bar{y}) = \exists x_1 \theta(\bar{y}, x_1)$. As $\theta$ is a shorter formula, by inductive hypothesis the claim has already been verified for $\theta$.

---

\(^1\)The notation in the preceding formula is, of course, a little sloppy, as the $b_i$ and $\bar{a}$ are not variables but elements of the structure over which we interpret. But we feel this notation improves readability.
Let \( \bar{a} \in N \) and assume \( i_0 < i_1 < \ldots < i_r \) are such that \( \bar{a} < b_{i_0} \), \( \varphi[\bar{a}] \) holds in \( N \) iff there exists a \( c \in N \) such that \( \theta[\bar{a}, c] \) holds in \( N \). Pick \( j_1 < j_2 < \ldots < j_r \) such that \( i_0 < j_1 \) and \( c < b_{j_1} \). By inductive hypothesis,

\[
N \models \theta[\bar{a}, c] \quad \text{iff} \quad M \models \theta^*[\bar{a}, c, b_{j_2}, \ldots, b_{j_r}].
\]

If we write it out, the expression on the right is

\[
M \models \forall x_2 < b_{j_2} \ldots Q x_r < b_{j_r} \, \psi(\bar{a}, c, x_2, \ldots, x_r).
\]

By choice of \( b_1 \), this is equivalent to

\[
M \models \exists x_1 < b_{j_1} \, \forall x_2 < b_{j_2} \ldots Q x_r < b_{j_r} \, \psi(\bar{a}, x_1, \ldots, x_r),
\]

in other words, it is equivalent to

\[
M \models \varphi^*[\bar{a}, b_{j_1}, \ldots, b_{j_r}].
\]

As \( i_0 < j_1 \) and the \((b_i)\) are diagonal indiscernibles for all \( \Delta_0 \) formulas in \( M \), the last expression is equivalent to

\[
M \models \varphi^*[\bar{a}, b_{i_0}, \ldots, b_{i_r}],
\]

which proofs the claim.

We can finally show that \( N \) satisfies induction. Recall that (Ind) is equivalent to the least number principle (LNP), as we saw in Section ???. Suppose \( N \models \varphi[a, \bar{c}] \), where \( \varphi(v, \bar{w}) \) is given in prenex normal form as

\[
\exists x_1 \forall x_2 \ldots Q x_n \, \psi(v, \bar{w}, \bar{x}), \quad \text{with } \psi \text{ quantifier free.}
\]

As before, we choose \( i_0 \) such that \( a, \bar{c} < b_{i_0} \). We can apply property (0.3) established in the Claim above and obtain the equivalence

\[
N \models \varphi[a, \bar{c}] \quad \text{iff} \quad M \models \exists x_1 < b_{i_0+1} \forall x_2 < b_{i_0+2} \ldots Q x_n < b_{i_0+n} \, \psi(a, \bar{c}, \bar{x}).
\]

Since induction (and hence LNP) holds in \( M \), there exists a least \( \hat{a} < b_{i_0} \) such that

\[
M \models \exists x_1 < b_{i_0+1} \forall x_2 < b_{i_0+2} \ldots Q x_n < b_{i_0+n} \, \psi(\hat{a}, \bar{c}, \bar{x}).
\]

By the definition of \( N \), the existence of \( \hat{a} \in N \), and the equivalence above, it follows that \( N \models \varphi[\hat{a}, \bar{c}] \). Finally, \( \hat{a} \) has to be the smallest witness to \( \varphi \) in \( N \), because any smaller witness would also be a smaller witness in \( M \). This concludes the proof of Proposition 4.46.