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

$$\exists x_1 \, \forall x_2 \, \ldots \, Q x_r \; \psi(\vec{y}, x_1, \ldots, x_r) \tag{0.1}$$

or

$$\forall x_1 \, \exists x_2 \, \ldots \, Q x_r \; \psi(\vec{y}, x_1, \ldots, x_r), \tag{0.2}$$

where Q is either $\exists$ or $\forall$, and ψ 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 $\varphi(\vec{y})$ in prenex normal form (0.1) we associate a Δ_0 formula $\varphi^*(\vec{y}, z_1, \ldots, z_r)$ given as

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

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

$$\mathcal{N} \vDash \varphi[\vec{a}] \;\Leftrightarrow\; \mathcal{M} \vDash \varphi[\vec{a}, b_{i_1}, \ldots, b_{i_r}]. \tag{0.3}$$

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

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

holds in $\mathcal{M}$.[1]

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

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

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

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

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

[1] The notation in the preceding formula is, of course, a little sloppy, as the b_i and $\vec{a}$ are not variables but elements of the structure over which we interpret. But we feel this notation improves readability.

Let $\vec{a} \in N$ and assume $i_0 < i_1 < \ldots < i_r$ are such that $\vec{a} < b_{i_0}$. $\varphi[\vec{a}]$ holds in $\mathcal{N}$ iff there exists a $c \in N$ such that $\theta[\vec{a}, c]$ holds in $\mathcal{N}$. Pick $j_1 < j_2 < \ldots < j_r$ such that $i_0 < j_1$ and $c < b_{j_1}$. By inductive hypothesis,

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

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

$$\mathcal{M} \models \forall x_2 < b_{j_2} \ldots Qx_r < b_{j_r}\ \psi(\vec{a}, c, x_2, \ldots, x_r).$$

By choice of b_1, this is equivalent to

$$\mathcal{M} \models \exists x_1 < b_{j_1}\ \forall x_2 < b_{j_2} \ldots Qx_r < b_{j_r}\ \psi(\vec{a}, x_1, \ldots, x_r),$$

in other words, it is equivalent to

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

As $i_0 < j_1$ and the (b_i) are diagonal indiscernibles for all Δ_0 formulas in $\mathcal{M}$, the last expression is equivalent to

$$\mathcal{M} \models \varphi^*[\vec{a}, b_{i_1}, \ldots, b_{i_r}],$$

which proofs the claim.

We can finally show that $\mathcal{N}$ satisfies induction. Recall that (Ind) is equivalent to the *least number principle* (LNP), as we saw in Section 4.1. Suppose $\mathcal{N} \models \varphi[a, \vec{c}]$, where $\varphi(v, \vec{w})$ is given in prenex normal form as

$$\exists x_1\ \forall x_2 \ldots Qx_n\ \psi(v, \vec{w}, \vec{x}), \qquad \text{with } \psi \text{ quantifier free.}$$

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

$$\mathcal{N} \models \varphi[a, \vec{c}] \ \text{ iff } \ \mathcal{M} \models \exists x_1 < b_{i_0+1}\ \forall x_2 < b_{i_0+2} \ldots Qx_n < b_{i_0+n}\ \psi(a, \vec{c}, \vec{x}).$$

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

$$\mathcal{M} \models \exists x_1 < b_{i_0+1}\ \forall x_2 < b_{i_0+2} \ldots Qx_n < b_{i_0+n}\ \psi(\hat{a}, \vec{c}, \vec{x}).$$

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