Fundamental Theorem of Algebra

Here we will use induction in the proof of the fundamental theorem of algebra to illustrate how induction is sometimes used in larger problems.

Definitions: A function $f : \mathbb{R} \to \mathbb{R}$ is a **polynomial** if it can be written in the form

$$f(x) = \sum_{i=0}^{d} c_i x^i = c_0 + c_1 x + c_2 x^2 + \cdots + c_d x^d,$$

where $c_i \in \mathbb{R}$ for $i = 0, 1, 2, \ldots, d$ are called the coefficients. If d is the exponent of the largest term that has a nonzero coefficients, we say the polynomial has **degree** d. A **zero**, or **root**, of the polynomial f is a number, a, such that $f(a) = 0$.

Examples:

- $f(x) = 5$ is a polynomial of degree 0 and it has zero real roots. (Note that the constant polynomial $f(x) = 0$ has degree undefined, not degree zero).
- $f(x) = x - 2$ is a polynomial of degree 1 and it has one real root $a = 2$.
- $f(x) = x^2 - 6x + 9 = (x - 3)^2$ is a polynomial of degree 2 and it has one real root, $a = 3$.
- $f(x) = x^3 - x = x(x^2 - 1)$ is a polynomial of degree 3 and it has three real roots, $a = 0, -1, +1$.

***Theorem*: (The Fundamental Theorem of Algebra)**
A polynomial of degree d has at most d real roots.

The proof below is based on two lemmas that are proved on the next page.

Proof: We use induction on d.

BASE STEP: If $d = 0$, then $f(x) = c_0$ for some nonzero constant c_0. Thus, $f(x)$ is never zero, so it has zero roots. Hence, in the $d = 0$ case the number of roots does not exceed d.

INDUCTIVE STEP: Assume every polynomial of degree k has at most k roots for some integer $k \geq 0$.

Let $f(x)$ be a polynomial of degree $k + 1$. We will show that $f(x)$ has at most $k + 1$ roots.
If $f(x)$ has no roots, then we are done, $0 \leq k + 1$.
If $f(x)$ has at least one root a, then, by Lemma 2, we can write $f(x) = (x - a)h(x)$ for some polynomial $h(x)$ with degree k. By the inductive hypothesis, $h(x)$ has at most k roots.
Since $x - a$ has one root and $h(x)$ has at most k roots, $f(x) = (x - a)h(x)$ has at most $k + 1$ roots.
Thus, in any case, $f(x)$ has at most $k + 1$ roots.
Hence, every polynomial of degree d has at most d roots. $\square$

Lemma 1: $\forall x, y \in \mathbb{R}$ and $\forall n \in \mathbb{N}$,

$$x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + xy^{n-2} + y^{n-1}).$$

Proof: We expand the right hand side using the distributive axiom to get

$$\begin{aligned}(x-y)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}) &= x(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}) \\ &\quad -y(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}) \\ &= x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + x^2y^{n-2} + xy^{n-1} \\ &\quad -x^{n-1}y - x^{n-2}y^2 - \cdots - x^2y^{n-2} - xy^{n-1} - y^n.\end{aligned}$$

Canceling all the middle terms, leaves only $x^n - y^n$. Thus, factoring in this way is always possible. □

Examples:
$x^2 - y^2 = (x-y)(x+y)$, $x^3 - y^3 = (x-y)(x^2 + xy + y^2)$, $x^4 - y^4 = (x-y)(x^3 + x^2y + xy^2 + y^3)$, *etc.*

Lemma 2: Suppose $f(x)$ is a polynomial of degree $d > 1$.
The number a is a zero of $f(x)$ if and only if $f(x) = (x-a)h(x)$ for some polynomial $h(x)$ of degree $d-1$.

Proof: We must prove both direction.

We prove the converse direction first. Assume $f(x) = (x-a)h(x)$ for some polynomial $h(x)$ of degree $d-1$. By substitution, $f(a) = (a-a)h(a) = 0 \cdot h(a) = 0$. Thus, $f(a) = 0$, so a is a zero of $f(x)$.

Now we prove for forward direction. Assume a is a real root of $f(x)$. Since $f(x)$ is of degree d, by definition, $f(x) = \sum_{i=0}^{d} c_i x^i$, with real number coefficients such that $c_d \neq 0$. Since a is a root of $f(x)$, $f(a) = 0$ and by substitution $\sum_{i=0}^{d} c_i a^i = 0$. By subtracting this expression (which is just subtracting zero), we can rewrite $f(x)$ as

$$f(x) = f(x) - 0 = f(x) - f(a) = \sum_{i=0}^{d} c_i x^i - \sum_{i=0}^{d} c_i a^i = \sum_{i=0}^{d} c_i (x^i - a^i).$$

The term corresponding to $i = 0$ cancels because $c_0(x^0 - a^0) = c_0(1-1) = 0$, so we have $f(x) = \sum_{i=1}^{d} c_i(x^i - a^i)$. By Lemma 1, for each $i > 0$, $x^i - a^i = (x-a)(x^{i-1} + x^{i-2}a + \cdots + xa^{i-2} + a^{i-1})$. By defining $h_i(x) = x^{i-1} + x^{i-2}a + \cdots + xa^{i-2} + a^{i-1}$, we now have $x^i - a^i = (x-a)h_i(x)$ where $h_i(x)$ is a polynomial of degree $i-1$. Hence, we can rewrite $f(x)$ as

$$\begin{aligned}f(x) &= \textstyle\sum_{i=1}^{d} c_i(x^i - a^i) \\ &= \textstyle\sum_{i=1}^{d} c_i(x-a)h_i(x) \\ &= (x-a)\textstyle\sum_{i=1}^{d} c_i h_i(x) \\ &= (x-a)h(x)\end{aligned}$$

Note that $h(x) = \sum_{i=1}^{d} c_i h_i(x) = \sum_{i=1}^{d} c_i(x^{i-1} + x^{i-2}a + \cdots + xa^{i-2} + a^{i-1})$, so $h(x)$ is a polynomial. And the term x^{d-1} occurs only once, when $i = d$, and it occurs with coefficient c_d which is not zero. Hence, $h(x)$ has degree $d-1$. □

Lemma 2 theorem effectively shows that we can always "factor out" the expression $(x-a)$ from a polynomial when a is a root.