17. Background

A nonempty bounded subset of $\mathbb{Z}$ contains a maximal element

This assertion is used several times in these notes; here is its proof.

Let m be an arbitrary element of the set M in question; there is at least one, by assumption. Further, let c be the bound. Then the algorithm

for `n=c:-1:m` **do: if** `n` $\in M$**, break, fi, od**

is guaranteed to halt after finitely many steps, and the current value of `n` is the maximal element of M.

Also, note the corollary that *a bounded function into the integers takes on its maximal value*: its range then contains a maximal element and any preimage of that maximal element will do.

Complex numbers

A complex number is of the form

$$z = a + \mathrm{i}b,$$

with a and b real numbers, called, respectively, the **real part of** z and the **imaginary part of** z, and i the **imaginary unit**, i.e.,

$$\mathrm{i} := \sqrt{-1}.$$

Actually, there are two complex numbers whose square is -1. We denote the other one by $-\mathrm{i}$. Be aware that, in parts of Engineering, the symbol j is used instead of i.

`MATLAB` works internally with (double precision) complex numbers. Both variables `i` and `j` in `MATLAB` are initialized to the value i. □

One adds complex numbers by adding separately their real and imaginary parts. One multiplies two complex numbers by multiplying out and rearranging, mindful of the fact that $\mathrm{i}^2 = -1$. Thus,

$$(a + \mathrm{i}b)(c + \mathrm{i}d) = ac + a\mathrm{i}d + b\mathrm{i}c - bd = (ac - bd) + \mathrm{i}(ad + bc).$$

Note that both addition and multiplication of complex numbers is commutative. Further, the product of $z = a + \mathrm{i}b$ with its **complex conjugate**

$$\overline{z} := a - \mathrm{i}b$$

is the nonnegative number

$$z\overline{z} = a^2 + b^2,$$

and its (nonnegative) squareroot is called the **absolute value** or **modulus** of z and denoted by

$$|z| := \sqrt{z\overline{z}}.$$

For $z \neq 0$, we have $|z| \neq 0$, hence $\overline{z}/|z|^2 = a/|z|^2 - \mathrm{i}b/|z|^2$ is a well-defined complex number. It is the **reciprocal** of z since $z\overline{z}/|z|^2 = 1$, of use for *division* by z. Note that, for any two complex numbers z and ζ,

$$|z\zeta| = |z||\zeta|.$$

It is very useful to visualize complex numbers as points in the so called **complex plane**, i.e., to identify the complex number $a + \mathrm{i}b$ with the point (a, b) in $\mathbb{R}^2$. With this identification, its absolute value corresponds to the (Euclidean) distance of the corresponding point from the origin.

The sum of two complex numbers corresponds to the vector sum of their corresponding points. The product of two complex numbers is most easily visualized in terms of the **polar form**

$$z = a + \mathrm{i}b = r\exp(\mathrm{i}\varphi),$$

with $r \geq 0$, hence $r = |z|$ its *modulus*, and $\varphi \in \mathbb{R}$ is called its **argument**. Indeed, for any real φ, $\exp(\mathrm{i}\varphi) = \cos(\varphi) + \mathrm{i}\sin(\varphi)$ has absolute value 1, and φ is the angle (in radians) that the vector (a, b) makes with the positive real axis. Note that, for $z \neq 0$, the argument, φ, is only defined up to a multiple of 2π, while, for $z = 0$, the argument is arbitrary. If now also $\zeta = \alpha + \mathrm{i}\beta = |\zeta|\exp(\mathrm{i}\psi)$, then, by the law of exponents,

$$z\zeta = |z|\exp(\mathrm{i}\varphi)|\zeta|\exp(\mathrm{i}\psi) = |z||\zeta|\exp(\mathrm{i}(\varphi + \psi)).$$

Thus, as already noted, the absolute value of the product is the product of the absolute values of the factors, while the argument of a product is the sum of the arguments of the factors.

For example, in as much as the argument of $\overline{z}$ is the negative of the argument of z, the argument of the product $z\overline{z}$ is necessarily 0. As another example, if $z = a + \mathrm{i}b$ is of modulus 1, then z lies on the unit circle in the complex plane, and so does any power z^k of z. In fact, then $z = \exp(\mathrm{i}\varphi)$ for some real number φ, and therefore $z^k = \exp(\mathrm{i}(k\varphi))$. Hence, the sequence $z^0, z^1, z^2, \cdots$ appears as a sequence of points on the unit circle, equally spaced around that circle, never accumulating anywhere unless $\varphi = 0$, i.e., unless $z = 1$.

(17.1) Lemma: Let z be a complex number of modulus 1. Then the sequence $z^0, z^1, z^2, \ldots$ of powers of z lies on the unit circle, but fails to converge except when $z = 1$.

Convergence of a scalar sequence

A subset Z of $\mathbb{C}$ is said to be **bounded** if it lies in some ball

$$B_r := \{z \in \mathbb{C} : |z| < r\}$$

of (finite) radius r. Equivalently, Z is bounded if, for some r, $|\zeta| < r$ for all $\zeta \in \mathrm{Z}$. In either case, the number r is called a **bound for** Z.

In particular, we say that the scalar sequence $(\zeta_1, \zeta_2, \ldots)$ is **bounded** if the set $\{\zeta_m : m \in \mathbb{N}\}$ is bounded. For example, the sequence $(1, 2, 3, \ldots)$ is not bounded.

(17.2) Lemma: The sequence $(\zeta^1, \zeta^2, \zeta^3, \cdots)$ is bounded if and only if $|\zeta| \leq 1$. Here, ζ^k denotes the kth power of the scalar z.

Proof: Assume that $|\zeta| > 1$. I claim that, for all m,

$$|\zeta^m| - 1 > (|\zeta| - 1)m. \tag{17.3}$$

This is certainly true for $m = 1$. Assume it correct for $m = k$. Then

$$|\zeta^{k+1}| - 1 = (|\zeta^{k+1}| - |\zeta^k|) + (|\zeta^k| - 1).$$

The first term on the right-hand side gives

$$|\zeta^{k+1}| - |\zeta^k| = (|\zeta| - 1)|\zeta|^{k-1} > |\zeta| - 1,$$

since $|\zeta| > 1$, while, for the second term, $|\zeta^k| - 1 > (|\zeta| - 1)k$ by induction hypothesis. Consequently,

$$|\zeta^{k+1}| - 1 > (|\zeta| - 1) + (|\zeta| - 1)k = (|\zeta| - 1)(k + 1),$$

showing that (17.3) also holds for $m = k + 1$.

In particular, for any given c, choosing m to be any natural number bigger than $c/(|\zeta| - 1)$, we have $|\zeta^m| > c$. We conclude that the sequence $(\zeta^1, \zeta^2, \zeta^3, \cdots)$ is unbounded when $|\zeta| > 1$.

Assume that $|\zeta| \leq 1$. Then, for any m, $|\zeta^m| = |\zeta|^m \leq 1^m = 1$, hence the sequence $(\zeta^1, \zeta^2, \zeta^3, \cdots)$ is not only bounded, it lies entirely in the **unit disk**

$$B_1^- := \{z \in \mathbb{C} : |z| \leq 1\}.$$

A sequence $(\zeta_1, \zeta_2, \zeta_3, \cdots)$ of (real or complex) scalars is said to **converge to the scalar** ζ, in symbols:

$$\lim_{m\to\infty} \zeta_m = \zeta,$$

if, for all $\varepsilon > 0$, there is some m_ε so that, for all $m > m_\varepsilon$, $|\zeta - \zeta_m| < \varepsilon$.

Assuming without loss the scalars to be complex, we can profitably visualize this definition as saying the following: Whatever small circle $\{z \in \mathbb{C} : |z - \zeta| = \varepsilon\}$ of radius ε we draw around the point ζ, *all* the terms of the sequence except the first few are inside that circle.

(17.4) Lemma: A convergent sequence is bounded.

Proof: If $\lim_{m\to\infty} \zeta_m = \zeta$, then there is some m_0 so that, for all $m > m_0$, $|\zeta - \zeta_m| < 1$. Therefore, for all m,

$$|\zeta_m| \leq r := |\zeta| + 1 + \max\{|\zeta_k| : k = 1{:}m_0\}.$$

Note that r is indeed a well-defined nonnegative number, since a *finite* set of real numbers always has a largest element. □

(17.5) Lemma: The sequence $(\zeta^1, \zeta^2, \zeta^3, \cdots)$ is convergent if and only if either $|\zeta| < 1$ or else $\zeta = 1$. In the former case, $\lim_{m\to\infty} \zeta^m = 0$, while in the latter case $\lim_{m\to\infty} \zeta^m = 1$.

Proof: Since the sequence is not even bounded when $|\zeta| > 1$, it cannot be convergent in that case. We already noted that it cannot be convergent when $|\zeta| = 1$ unless $\zeta = 1$, and in that case $\zeta^m = 1$ for all m, hence also $\lim_{m\to\infty} \zeta^m = 1$.

This leaves the case $|\zeta| < 1$. Then either $|\zeta| = 0$, in which case $\zeta^m = 0$ for all m, hence also $\lim_{m\to\infty} \zeta^m = 0$. Else, $0 < |\zeta| < 1$, therefore $1/\zeta$ is a well-defined complex number of modulus greater than one, hence, as we showed earlier, $1/|\zeta^m| = |(1/\zeta)^m|$ grows monotonely to infinity as $m \to \infty$. But this says that $|\zeta^m|$ decreases monotonely to 0. In other words, $\lim_{m\to\infty} \zeta^m = 0$. □

Horner, or: How to divide a polynomial by a linear factor

Recall that, given the polynomial p and one of its roots, μ, the polynomial $q := p/(\cdot - \mu)$ can be constructed by **synthetic division**. This process is also known as **nested multiplication** or **Horner's scheme** as it is used, more generally, to evaluate a polynomial efficiently. Here are the details, for a polynomial of degree ≤ 3.

If $p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$, and z is any scalar, then

$$p(z) = \underbrace{a_0 + z\,\underbrace{(a_1 + z\,\underbrace{(a_2 + z\,\underbrace{a_3}_{=:b_3})}_{=:b_2})}_{=a_1+zb_2=:b_1}}_{a_0+zb_1=:b_0}.$$

In other words, we write such a polynomial in **nested form** and then evaluate from the inside out. Each step is of the form

horner
$$b_j := a_j + z b_{j+1}, \tag{17.6}$$

it involves one multiplication and one addition. The last number calculated is b_0; it is the value of p at z. There are 3 such steps for our cubic polynomial (the definition $b_3 := a_3$ requires no calculation!). So, for a polynomial of degree n, we would use n multiplications and n additions.

Now, not only is b_0 of interest, since it equals $p(z)$; the other b_j are also useful since

$$p(t) = b_0 + (t - z)(b_1 + b_2 t + b_3 t^2).$$

We verify this by multiplying out and rearranging terms according to powers of t. This gives

$$\begin{array}{rcccccccc} b_0 + (t-z)(b_1 + b_2 t + b_3 t^2) & = & b_0 & + & b_1 t & + & b_2 t^2 & + & b_3 t^3 \\ & & -z b_1 & - & z b_2 t & - & z b_3 t^2 & & \\ & = & b_0 - z b_1 & + & (b_1 - z b_2) t & + & (b_2 - z b_3) t^2 & + & b_3 t^3 \\ & = & a_0 & + & a_1 t & + & a_2 t^2 & + & a_3 t^3 \end{array}$$

The last equality holds since, by (17.6),

$$b_j - z b_{j+1} = a_j$$

for $j < 3$ while $b_3 = a_3$ by definition.

(17.7) Nested Multiplication (aka Horner): To evaluate the polynomial $p(t) = a_0 + a_1 t + \cdots + a_k t^k$ at the point z, compute the sequence $(b_0, b_1, \ldots, b_k)$ by the prescription

$$b_j := \begin{cases} a_j & \text{if } j = k; \\ a_j + z b_{j+1} & \text{if } j < k. \end{cases}$$

Then $p(t) = b_0 + (t - z) q(t)$, with

$$q(t) := b_1 + b_2 t + \cdots + b_k t^{k-1}.$$

In particular, if z is a root of p (hence $b_0 = 0$), then

$$q(t) = p(t)/(t - z).$$

Since $p(t) = (t - z) q(t)$, it follows that $\deg q < \deg p$. This provides another proof (see (3.21)) for the *easy* part of the *Fundamental Theorem of Algebra*, namely that a polynomial of degree k has at most k roots.