Chứng minh số π là số siêu việt (Theo Ian Stewart, University of Warwick, 1970)

(*Các bổ đề, định lí sử dụng trong chứng minh được trình bày chi tiết trong phần cuối bài viết.*)

Cách chứng minh này chỉ sử dụng các kiến thức cơ bản của chương trình đại học, không sử dụng các chuyên đề sâu hơn về đại số. Chứng minh bằng phản chứng.

Giả sử π là số đại số, do vậy theo bổ đề 1, $i\pi$ cũng là số đại số. (*Ta có thể chứng minh trực tiếp mà không sử dụng bổ đề 1, bằng việc xét đa thức $g(x) \in \mathbb{Q}[x], g(\pi) = 0$. Khi đó $u(x) = g(ix) \cdot g(-ix) \in \mathbb{Q}[x]$ và $u(i\pi) = 0$).*

Gọi r là bậc của đa thức tối tiểu $g(x)$ của $i\pi$ và b là hệ số bậc cao nhất $g(x) = bx^r + b_1x^{r-1} + \cdots + b_{r-1}x + b_r$. Kí hiệu $\omega_1 = i\pi, \omega_2, ..., \omega_r$ là các nghiệm khác nhau của đa thức tối tiểu. (Chú ý rằng theo bổ đề 2, $b\omega_i$ là các số nguyên đại số). Theo công thức Euler $e^{\omega_1} = e^{i\pi} = -1$, ta có

$$(1 + e^{\omega_1})(1 + e^{\omega_2}) \cdots (1 + e^{\omega_r}) = 0. \quad (1)$$

Nhân tung biểu thức ở vế trái, ta được

$$1 + (e^{\omega_1} + e^{\omega_2} + \cdots + e^{\omega_r}) + (e^{\omega_1+\omega_2} + e^{\omega_1+\omega_3} + \cdots) + \cdots + (e^{\omega_1+\omega_2+\cdots+\omega_r}) = 0$$

Vế trái là một hàm đối xứng theo n biến, gồm 2^r số hạng, mỗi số hạng có dạng e^{Φ} mà

$$\Phi = \varepsilon_1\omega_1 + \varepsilon_2\omega_2 + \cdots + \varepsilon_r\omega_r, \ \varepsilon_i \in \{0, 1\}. \quad (*)$$

Giả sử có n giá trị $\Phi_1, \Phi_2, ...\Phi_n$ khác 0 trong số các Φ ở $(*)$ (và $q = 2^r - n$ giá trị còn lại đều bằng 0) khi đó đẳng thức (1) rút gọn thành

$$q + e^{\Phi_1} + e^{\Phi_2} + \cdots + e^{\Phi_n} = 0.$$

Với p là số nguyên tố đủ lớn sẽ xác định sau, xét đa thức bậc $m = np + p - 1$

$$f(x) = b^{np}x^{p-1}(x - \Phi_1)^p(x - \Phi_2)^p \cdots (x - \Phi_n)^p.$$

Sử dụng định lí cơ bản của đa thức đối xứng ta sẽ chỉ ra $f(x) \in \mathbb{Z}[x]$. Thật vậy với n giá trị Φ_i khác 0 nên $\prod_{i=1}^{2^r}(x - \Phi_i) = x^{2^r-n}\prod_{i=1}^{n}(x - \Phi_i)$ là hàm đối xứng theo các biến $\omega_1, \omega_2, ..., \omega_r$. Do đó các hệ số của $f(x)$ là các số hữu tỉ. Mặt khác $f(x)$ chứa các thừa số $\prod_{i=1}^{n}(bx - b\Phi_i)$ mà khi nhân bung ra rồi rút gọn, ta đã thực hiện các phép cộng, trừ, nhân các số nguyên đại số $b\Phi_i$ với nhau (bổ đề 2). Vậy các hệ số của $f(x)$ là các số nguyên đại số. Áp dụng bổ đề 3 suy ra $f(x) \in \mathbb{Z}[x]$. Định nghĩa $J = I(\Phi_1) + I(\Phi_2) + \cdots + I(\Phi_n)$ và sử dụng bổ đề 4,

$$J = \sum_{k=1}^{n} e^{\Phi_k} \sum_{j=0}^{m} f^{(j)}(0) - \sum_{k=1}^{n}\sum_{j=0}^{m} f^{(j)}(\Phi_k) = \sum_{j=0}^{m} f^{(j)}(0) \sum_{k=1}^{n} e^{\Phi_k} - \sum_{j=0}^{m}\sum_{k=1}^{n} f^{(j)}(\Phi_k) = -q\sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m}\sum_{k=1}^{n} f^{(j)}(\Phi_k).$$

Các hàm $f^{(j)}(x)$ cũng là các đa thức hệ số nguyên đối xứng theo các biến $b\Phi_1, b\Phi_2, ..., b\Phi_n$, do đó cũng là các đa thức hệ số nguyên đối xứng theo các biến $b\omega_1, b\omega_2, ..., b\omega_n$ và theo định lí cơ bản của đa thức đối xứng $\sum_{k=1}^{n} f^{(j)}(\Phi_k)$ là số hữu tỉ. Áp dụng bổ đề 2 và 3 suy ra nó là số nguyên.

Các số $\Phi_1, \Phi_2, ...\Phi_n$ là các nghiệm bậc p của f và số 0 là nghiệm bậc $p-1$ của f, hay $f^{(j)}(\Phi_k) = 0$ với mọi $j < p$, do vậy chỉ số j dưới dấu $\sum$ trong J chỉ xuất phát từ $j = p - 1$ trong số hạng đầu và xuất phát từ $j = p$ trong số hạng thứ hai

$$J = -q \sum_{j=p-1}^{m} f^{(j)}(0) - \sum_{j=p}^{m}\sum_{k=1}^{n} f^{(j)}(\Phi_k). \quad (2)$$

Từ định nghĩa hàm f, ta thấy các số hạng $f^{(j)}(\Phi_k)$ trong tổng kép, với $j \geq p$ đều chứa thừa số $p!$, vì vậy nó chia hết cho $p!$ và sử dụng công thức Leibnitz tính đạo hàm cấp cao, $f^{(j)}(0)$ cũng chia hết cho $p!$ với $j \geq p$. Ta lại có (vẫn sử dụng công thức Leibnitz)

$$f^{(p-1)}(0) = b^{np}(-1)^{np}(p-1)!(\Phi_1\Phi_2\cdots\Phi_n)^p.$$

Như vậy $f^{(p-1)}(0)$ không chia hết cho $p!$ với p là số nguyên tố đủ lớn. Từ hệ thức (2) suy ra J chia hết cho $(p-1)!$ và không chia hết cho $p!$. Do đó $J \neq 0$ và

$$|J| \geq (p-1)! \quad (3)$$

Mặt khác từ định nghĩa của $I(\Phi_i) = \displaystyle\int_0^{\Phi_i} e^{\Phi_i - u} f(u)\, du$, trong đó

$$f(x) = b^{np}x^{p-1}(x-\Phi_1)^p(x-\Phi_2)^p\cdots(x-\Phi_n)^p$$

ta có ước lượng khá thô $|f(x)| \leq (|b|^n)^p \cdot (\max|\Phi_i|)^{p-1} \cdot A^p \leq C \cdot B^p$, trong đó các hằng số B, C không phụ thuộc vào p. Suy ra

$$|J| \leq \sum_{k=1}^{n} |I(\Phi_k)| \leq \sum_{k=1}^{n} |\Phi_k| \cdot e^{|\Phi_k|} \cdot C \cdot B^p = D \cdot B^p, \quad (4)$$

Với $D = \sum_{k=1}^{n} |\Phi_k| \cdot e^{|\Phi_k|}$ không phụ thuộc vào p. Kết hợp (3) và (4), sử dụng bất đẳng thức cơ bản $e^p \geq \dfrac{p^{p-1}}{(p-1)!}$ ta được

$$p^{p-1}e^{-p} \leq (p-1)! \leq D \cdot B^p, \qquad \text{vô lí khi số } p \text{ đủ lớn.}$$

Bằng cách đó ta đã hoàn thành việc chứng minh π là số siêu việt.

Các bổ đề, định lí sử dụng trong chứng minh trên

Định lí cơ bản của đa thức đối xứng. *Mọi đa thức đối xứng hệ số nguyên $f(x_1, x_2, ..., x_n) \in \mathbb{Z}[x_1, x_2, ..., x_n]$ luôn biểu diễn duy nhất thành một đa thức cũng hệ số nguyên $F(s_1, s_2, ..., s_n)$ mà các biến của F là các hàm đa thức cơ bản đối xứng.* Các đa thức cơ bản đối xứng $s_1, s_2, ..., s_n$ được định nghĩa

$$s_1 = x_1 + x_2 + \cdots + x_n,$$

$$s_2 = x_1x_2 + x_1x_3 + \cdots + x_ix_j + \cdots + x_{n-1}x_n,$$

$$\cdots \quad \cdots \quad \cdots$$

$$s_k = x_1x_2...x_k + \cdots + x_{i_1}x_{i_2}...x_{i_k} + \cdots,$$

$$\cdots \quad \cdots \quad \cdots$$

$$s_n = x_1x_2...x_n.$$

Ví dụ 1. Đa thức đối xứng 2 biến $f(x,y) = 2x^2 - 3xy + 2y^2 = 2(x+y)^2 - 7xy = 2s_1^2 - 7s_2$.

Ví dụ 2. Đa thức đối xứng 2 biến $g(x,y) = 2x^3 + 2y^3 = 2(x+y)^3 - 6xy(x+y) = 2s_1^3 - 6s_1s_2$.

Ví dụ 3. Đa thức đối xứng 3 biến

$$h(x,y,z) = 2x^3 + 2y^3 + 2z^3 = 6xyz + 2(x+y+z)^3 - 6(x+y+z)(xy+xz+yz) = 2s_1^3 - 6s_1s_2 + 6s_3.$$

Định lí Viète. *Cho đa thức $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \in \mathbb{C}[x]$ bậc n. Kí hiệu $x_1, x_2, ..., x_n$ là các nghiệm của f. (Lưu ý các nghiệm bội k được viết lặp lại k lần). Khi đó*

$$s_k = x_1x_2...x_k + \cdots + x_{i_1}x_{i_2}...x_{i_k} + \cdots = (-1)^{n-k}\frac{a_k}{a_0} \quad \text{với mọi} \quad 1 \le k \le n.$$

Định nghĩa. Một số phức $\alpha \in \mathbb{C}$ được gọi là số đại số nếu nó là nghiệm của một đa thức hệ số hữu tỉ. Tập các số đại số được kí hiệu $\overline{\mathbb{Q}}$. Một số phức $\alpha \in \mathbb{C}$ được gọi là số nguyên đại số nếu nó là nghiệm của đa thức

$$x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n,$$

trong đó các hệ số $a_i \in \mathbb{Z}$ là các số nguyên. Tập các số nguyên đại số được kí hiệu $\overline{\mathbb{Z}}$.

Bổ đề 1. *Nếu α, β là các số đại số thì $\alpha \pm \beta, \alpha \cdot \beta$, và $\frac{\alpha}{\beta}, (\beta \neq 0)$ cũng là các số đại số. Hơn nữa nếu α, β là các số nguyên đại số thì $\alpha \pm \beta, \alpha \cdot \beta$ cũng là các số nguyên đại số.*

Bổ đề chứng tỏ $\overline{\mathbb{Q}}$ là một trường và $\overline{\mathbb{Z}}$ là một vành. Phương pháp chứng minh, chỉ sử dụng không gian tuyến tính là một ví dụ hay về vẻ đẹp của toán học. Đó là lí do để tôi giới thiệu nó ở đây.

Chứng minh. Ta sẽ chứng minh $\overline{\mathbb{Q}}$ là một trường, nửa sau của bổ đề ($\overline{\mathbb{Z}}$ là một vành) được chứng minh cũng dựa theo ý tưởng như vậy. Ta không viết chi tiết ra đây.

Số phức $\alpha \in \mathbb{C}$ là số đại số khi và chỉ khi không gian véc tơ $U = \mathscr{L}(1, \alpha, \alpha^2, ...)$ trên trường số hữu tỉ $\mathbb{Q}$ là không gian hữu hạn chiều. Giả sử $\dim U = d$ khi đó $1, \alpha, \alpha^2, ..., \alpha^d$ phụ thuộc tuyến tính với các hệ số thuộc $\mathbb{Q}$ và do đó $\alpha \in \overline{\mathbb{Q}}$. Ngược lại, nếu

$$\alpha^n + a_1\alpha^{n-1} + \cdots + a_n = 0 \quad \Rightarrow \quad \alpha^n = -a_1\alpha^{n-1} + \cdots + a_n \in \mathscr{L}(1, \alpha, ..., \alpha^{n-1})$$

Suy ra $\alpha^{n+1} = -a_1\alpha^n + \cdots + a_n\alpha \in \mathscr{L}(1, \alpha, ..., \alpha^{n-1})$ và cứ thế $\alpha^{n+1}, \alpha^{n+2}, ... \in \mathscr{L}(1, \alpha, ..., \alpha^{n-1})$.

Bây giờ giả sử $\alpha, \beta \in \overline{\mathbb{Q}}$ và $U = \mathscr{L}(1, \alpha, \alpha^2, ...)$, $V = \mathscr{L}(1, \beta, \beta^2, ...)$ là các không gian hữu hạn chiều. Khi đó

$$\alpha U \subset U, \ \beta V \subset V.$$

Kí hiệu

$$UV = \mathscr{L}(uv, u \in U, v \in V) = \{u_1v_1 + u_2v_2 + \cdots + u_rv_r, \ r \ge 1, \ u_i \in U, v_i \in V\}$$

cũng là không gian hữu hạn chiều và

$$(\alpha + \beta)UV \subset UV, \quad (\alpha \cdot \beta)UV \subset UV.$$

Vậy $\alpha + \beta, \alpha \cdot \beta \in \overline{\mathbb{Q}}$.

Cuối cùng ta chỉ cần chứng minh nếu $\alpha \in \overline{\mathbb{Q}}, \alpha \neq 0$ thì $1/\alpha \in \overline{\mathbb{Q}}$. Thậy vậy giả sử α là nghiệm của đa thức

$$f(x) = b_0 x^r + b_1 x^{r-1} + \cdots + b_{r-1} x + b_r$$

thì $1/\alpha$ là nghiệm của đa thức

$$x^r f(1/x) = b_r x^r + b_{r-1} x^{r-1} + \cdots + b_1 x + b_0, \quad \text{đ.p.c.m.}$$

Bổ đề 2. *Nếu α là số đại số và $g(x) \in \mathbb{Z}[x]$ là đa thức tối tiểu của nó với b là hệ số bậc cao nhất. Khi đó $b\alpha$ là số nguyên đại số.*

Chứng minh. Giả sử α là nghiệm của đa thức

$$f(x) = bx^r + b_1 x^{r-1} + \cdots + b_{r-1} x + b_r,$$

khi đó $b\alpha$ thỏa mãn phương trình

$$x^r + b_1 x^{r-1} + b b_2 x^{r-2} + \cdots + b^{r-2} b_{r-1} x + b^{r-1} b_r = 0,$$

Bổ đề 3. *Nếu α là số nguyên đại số đồng thời cũng là cũng là số hữu tỉ thì α là số nguyên thông thường ($\in \mathbb{Z}$).*

Chứng minh. Giả sử $\alpha = \frac{r}{s}$ thương của 2 số nguyên (r và s nguyên tố cùng nhau) và là nghiệm của đa thức

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n = 0.$$

Khi đó

$$r^n + a_1 r^{n-1} s + \cdots + a_n s^n = 0,$$

suy ra s là ước số r^n. Mặt khác r và s nguyên tố cùng nhau nên $s = \pm 1$ hay $\alpha = \frac{r}{s} \in \mathbb{Z}$, đ.p.c.m.

Bổ đề 4. Với f là đa thức bậc m và với mọi $t \in \mathbb{C}$, tích phân

$$I(t, f) = \int_0^t e^{t-u} f(u)\, du = e^t \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(t).$$

Bằng phương pháp tích phân từng phần ta có

$$I(t, f) = e^t f(0) - f(t) + I(t, f').$$

Tiếp tục tính tích phân từng phần $I(t, f')$

$$I(t, f) = e^t f(0) - f(t) + I(t, f') = e^t\big(f(0) + f'(0)\big) - \big(f(t) + f'(t)\big) + I(t, f'').$$

Bằng quy nạp ta dễ dàng chứng minh bổ đề trên.