Divison of Complex Numbers, Square Roots, Higher-Order Roots, Multiple-Angle Formulae, and Trigonometric Sums

(1) To divide $\alpha+i\beta$ by $\gamma+i\delta$, we make the denominator root by multiplying both the numerator and the denominator by the complex conjugate of the denominator.

$$\frac{\alpha+i\beta}{\gamma+i\delta}=\frac{(\alpha+i\beta)(\gamma-i\delta)}{(\gamma+i\delta)(\gamma-i\delta)}=\frac{(\alpha\gamma+\beta\delta)+i(\beta\gamma-\alpha\delta)}{\gamma^2+\delta^2}.$$

(2) To get the square root of $x+iy$, we let its answer be $\alpha+i\beta$ so that we have the equations

$$\begin{cases} x^2-y^2=\alpha \\ 2xy=\beta \end{cases}$$

and, when we square the first equation and add to its the square of the second equation, we get

$$\left(x^2+y^2\right)^2=\left(x^2-y^2\right)^2+4x^2y^2=\alpha^2+\beta^2.$$

(The motivation for this step is that the explicit formula for the solution of a quadratic equation enables us to solve for two unknowns when both their sum and their product are known. The formula $(a+b)^2=(a-b)^2+4ab$ applied to the case of $a=x^2$ and $b=y^2$ enables us to solve for two unknowns when both their *difference* and their product are known.) Hence

$$x^2+y^2=\sqrt{\alpha^2+\beta^2}$$

and, when we couple with the first equation $x^2-y^2=\alpha$, we get

$$\begin{cases} x^2=\frac{1}{2}\left(\alpha+\sqrt{\alpha^2+\beta^2}\right) \\ y^2=\frac{1}{2}\left(-\alpha+\sqrt{\alpha^2+\beta^2}\right) \end{cases}.$$

If we just take the square roots here, we end up with four possibilities for x and y instead of the two which we know we should have. In order to rule out two, we go back to the equation $2xy=\beta$ whose square we have been using instead, thereby resulting in less precise conclusions. We write

$$\begin{cases} x=\epsilon_1\sqrt{\frac{1}{2}\left(\alpha+\sqrt{\alpha^2+\beta^2}\right)} \\ y=\epsilon_2\sqrt{\frac{1}{2}\left(-\alpha+\sqrt{\alpha^2+\beta^2}\right)} \end{cases},$$

where $\epsilon_1 = \pm 1$ and $\epsilon_2 = \pm 1$. From

$$\beta = 2xy = 2\left(\epsilon_1\sqrt{\frac{1}{2}\left(\alpha+\sqrt{\alpha^2+\beta^2}\right)}\right)\left(\epsilon_2\sqrt{\frac{1}{2}\left(-\alpha+\sqrt{\alpha^2+\beta^2}\right)}\right)$$
$$= \epsilon_1\epsilon_2\sqrt{\left(\sqrt{\alpha^2+\beta^2}\right)-\alpha^2} = \epsilon_1\epsilon_2|\beta|,$$

we conclude that $\epsilon_2 = \epsilon_1\frac{\beta}{|\beta|}$. We can choose $\epsilon_1 = \pm 1$ and get the final answer

$$\sqrt{\alpha+i\beta} = \pm\left(\sqrt{\frac{\alpha+\sqrt{\alpha^2+\beta^2}}{2}} + i\frac{\beta}{|\beta|}\sqrt{\frac{-\alpha+\sqrt{\alpha^2+\beta^2}}{2}}\right).$$

(3) *de Moivre's Formula*

$$(\cos\varphi + i\sin\varphi)^n = \cos(n\varphi) + i\sin(n\varphi).$$

This comes from the definition of $e^{ix} = \cos x + i\sin x$ (which is the *Euler formula*) and the verification that

$$(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) = \cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2)$$

resulting from

$$\begin{cases} \cos(\varphi_1+\varphi_2) = \cos\varphi_1\,\cos\varphi_2 - \sin\varphi_1\,\sin\varphi_2 \\ \sin(\varphi_1+\varphi_2) = \sin\varphi_1\,\cos\varphi_2 + \cos\varphi_1\,\sin\varphi_2. \end{cases}$$

The de Moivre formula gives the roots $z = \rho(\cos\theta + i\sin\theta)$ of order n for the complex number $a = r(\cos\varphi + i\sin\varphi)$ as

$$z = \sqrt[n]{r}\left(\cos\left(\frac{\varphi}{n} + k\frac{2\pi}{n}\right) + i\sin\left(\frac{\varphi}{n} + k\frac{2\pi}{n}\right)\right)$$

for $k = 0, 1, \cdots, n-1$. By expanding the right-hand side of

$$(\cos\varphi + i\sin\varphi)^n = \cos(n\varphi) + i\sin(n\varphi).$$

in terms of binomial coefficients and equating the real part and the imaginary part of both sides, we can express $\cos(n\varphi)$ and $\sin(n\varphi)$ as explicit polynomials of degree $\leq n$ in terms of $\cos\varphi$ and $\sin\varphi$ with rational numbers as coefficients. Such expressions are the *multiple-angle formulas*. More precisely, by equating the real and imaginary parts of

$$\sum_{k=0}^{n} i^k \binom{n}{k} \cos^{n-k}\varphi \sin^k\varphi = \cos(n\varphi) + i\sin(n\varphi),$$

we obtain

$$\cos(n\varphi) = \sum_{\substack{0\leq k\leq n \\ k \text{ even}}} (-1)^{\frac{k}{2}} \binom{n}{k} \cos^{n-k}\varphi \sin^k\varphi$$

and

$$\sin(n\varphi) = \sum_{\substack{0\leq k\leq n \\ k \text{ odd}}} (-1)^{\frac{k-1}{2}} \binom{n}{k} \cos^{n-k}\varphi \sin^k\varphi.$$

(4) By using the Euler formula, we can use the summation of geometric series of e^{ix} to sum certain trigonometric series. Important examples are the Dirichlet kernel $D_n(x)$ and the Féjer kernel $F_n(x)$:

$$D_n(x) = \frac{1}{2\pi}\sum_{k=-n}^{n} e^{ikx} = \frac{1}{2\pi}\frac{\sin\left(n+\frac{1}{2}\right)x}{\sin\frac{x}{2}},$$

$$F_n(x) = \frac{D_0(x) + D_1(x) + \cdots + D_{n-1}(x)}{n} = \frac{1}{2\pi n}\frac{\sin^2\frac{nx}{2}}{\sin^2\frac{x}{2}}.$$

The motivation for the Dirichlet kernel $D_n(x)$ comes from the convergent question of the Fourier series $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ of a function $f(x)$ on $\mathbb{R}$ with periodicity 2π (which means that $f(x+2\pi) = f(x)$), where the n-th *Fourier coefficient* c_n of $f(x)$ is defined by

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}dx.$$

Later in this course we will discuss in detail the theory of Fourier series. At this point we just use it as a motivation for an example to illustrate the method of summing certain geometric series by using the Euler formula and the geometric series. To consider the question of convergence to $f(x)$ of the Fourier series $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ of $f(x)$, we want to express the Fourier series $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ directly in terms of $f(x)$ by replacing the Fourier coefficient c_n by the expression in its definition to obtain

$$\begin{aligned}\sum_{n=-\infty}^{\infty} c_n e^{inx} &= \lim_{n\to\infty} \sum_{k=-n}^{n} \left(\frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt \right) e^{inx} \\ &= \lim_{n\to\infty} \int_0^{2\pi} f(t) \left(\frac{1}{2\pi} \sum_{k=-n}^{n} e^{in(x-t)} \right) dt \\ &= \lim_{n\to\infty} \int_0^{2\pi} f(x-t) \left(\frac{1}{2\pi} \sum_{k=-n}^{n} e^{int} \right) dt \\ &= \lim_{n\to\infty} \int_0^{2\pi} f(x-t) D_n(t) dt.\end{aligned}$$

This is the motivation for the Dirichlet kernel $D_n(x)$. The verification of the closed-form expression of $D_n(x)$ (from summing the series) is as follows.

$$\begin{aligned}\sum_{k=-n}^{n} e^{ikx} &= \sum_{k=0}^{n} \left(e^{ix}\right)^k + \sum_{k=-n}^{-1} \left(e^{ix}\right)^k \\ &= \sum_{k=0}^{n} \left(e^{ix}\right)^k + e^{-inx} \sum_{k=0}^{n-1} \left(e^{ix}\right)^k \\ &= \frac{1-\left(e^{ix}\right)^{n+1}}{1-e^{ix}} + e^{-inx} \frac{\left(e^{ix}\right)^{-n} - 1}{1-e^{ix}} \\ &= \frac{1-\left(e^{ix}\right)^{n+1}}{1-e^{ix}} + \frac{1-\left(e^{ix}\right)^{n}}{1-e^{ix}} \\ &= \frac{\left(e^{ix}\right)^{-n} - \left(e^{ix}\right)^{n+1}}{1-e^{ix}} = \frac{\left(e^{ix}\right)^{-n-\frac{1}{2}} - \left(e^{ix}\right)^{n+\frac{1}{2}}}{\left(e^{ix}\right)^{-\frac{1}{2}} - \left(e^{ix}\right)^{\frac{1}{2}}} \\ &= \frac{\sin\left(n+\frac{1}{2}\right)x}{\sin\frac{x}{2}}.\end{aligned}$$

The key is take away $\left(e^{ix}\right)^{\frac{1}{2}}$ (whose power is the mid-point of the powers of the two terms 1 and e^{ix} whose difference forms the denominator) from both the numerator and the denominator simultaneously so that the denominator becomes $\left(e^{ix}\right)^{-\frac{1}{2}} - \left(e^{ix}\right)^{\frac{1}{2}}$ which is purely imaginary. Note that the technique of dividing two complex numbers by multiplying both the numerator and the denominator by the complex conjugate of the latter will not be able to yield as easily the desired final expression.

(5) *Cauchy-Schwarz Inequality.*

$$\left|\sum_{j=1}^{n} a_j\overline{b_j}\right|^2 \leq \left(\sum_{j=1}^{n}|a_j|^2\right)\left(\sum_{j=1}^{n}|b_j|^2\right),$$

where the "equal" sign holds if and only if there exist $\lambda, \mu \in \mathbb{C}$ not both zero such that $\lambda a_j + \mu b_j = 0$ for all $1 \leq j \leq n$. We offer two proofs. The first proof is by Lagrange's identity which can be regarded as a generalization of the identity $1 = \cos^2\theta + \sin^2\theta$ for $\theta \in \mathbb{R}$. Lagrange's identity is

$$\left(\sum_{j=1}^{n}|a_j|^2\right)\left(\sum_{k=1}^{n}|b_k|^2\right) = \left|\sum_{j=1}^{n} a_j\overline{b_j}\right|^2 + \sum_{1\leq j<k\leq n}|a_jb_k - a_kb_j|^2 .$$

It can be regarded as a generalization of the identity $1 = \cos^2\theta + \sin^2\theta$ for $\theta \in \mathbb{R}$, because when $n = 3$ and when all a_j and b_k are real, we can rewrite it as

$$\|\vec{a}\|^2 \left\|\vec{b}\right\|^2 = \left\|\vec{a}\cdot\vec{b}\right\|^2 + \left\|\vec{a}\times\vec{b}\right\|^2$$

which is the same as the product of $\|\vec{a}\|^2 \left\|\vec{b}\right\|^2$ with $1 = \cos^2\theta + \sin^2\theta$, where θ is the angle between the two vectors $\vec{a} = (a_1, a_2, a_3)$ and $\vec{b} = (b_1, b_2, b_3)$. Lagrange's identity implies immediately the Cauchy-Schwarz inequality with equality holding if and only if $a_jb_k - a_kb_j = 0$ for all $1 \leq j < k \leq n$, which means that if some $b_{j_0} \neq 0$ then $a_k + \frac{-a_{j_0}}{b_{j_0}}b_k = 0$ and we can set $\lambda = 1$ and $\mu = \frac{-a_{j_0}}{b_{j_0}}$ and if some $a_{j_0} \neq 0$ then $b_k + \frac{-b_{j_0}}{a_{j_0}}a_k = 0$ and we can set $\lambda = \frac{-b_{j_0}}{a_{j_0}}$ and $\mu = 1$. The verification of Lagrange's identity is as follows.

$$\sum_{1\leq j<k\leq n}|a_jb_k - a_kb_j|^2 = \frac{1}{2}\sum_{1\leq j,k\leq n}|a_jb_k - a_kb_j|^2$$

$$= \frac{1}{2}\sum_{1\leq j,k\leq n}(a_jb_k - a_kb_j)\overline{(a_jb_k - a_kb_j)}$$

$$\sum_{1\leq j,k\leq n}\left(a_jb_k\overline{a_j}\overline{b_k}-a_jb_k\overline{a_k}\overline{b_j}-a_kb_j\overline{a_j}\overline{b_k}+a_kb_j\overline{a_k}\overline{b_j}\right)$$

$$=\sum_{1\leq j,k\leq n}\left(a_jb_k\overline{a_j}\overline{b_k}-a_jb_k\overline{a_k}\overline{b_j}\right)$$

$$=\left(\sum_{j=1}^n a_j\overline{a_j}\right)\left(\sum_{k=1}^n b_k\overline{b_k}\right)-\left(\sum_{j=1}^n a_j\overline{b_j}\right)\left(\sum_{k=1}^n a_k\overline{b_k}\right)$$

$$=\left(\sum_{j=1}^n |a_j|^2\right)\left(\sum_{k=1}^n |b_k|^2\right)-\left|\sum_{j=1}^n a_j\overline{b_j}\right|^2,$$

because interchanging the two indices j and k yields

$$\sum_{1\leq j,k\leq n} a_jb_k\overline{a_k}\overline{b_j}=\sum_{1\leq j,k\leq n} a_kb_j\overline{a_j}\overline{b_k},$$
$$\sum_{1\leq j,k\leq n} a_jb_k\overline{a_j}\overline{b_k}=\sum_{1\leq j,k\leq n} a_kb_j\overline{a_k}\overline{b_j}.$$

The second proof uses the orthogonal projection $\lambda\vec{b}$ of one vector $\vec{a}=(a_1,\cdots,a_n)\in\mathbb{C}^n$ onto another vector $\vec{b}=(b_1,\cdots,b_n)\in\mathbb{C}^n$ when $\lambda\in\mathbb{C}$ is chosen as

$$\frac{\left\langle\vec{a},\,\vec{b}\right\rangle}{\|\vec{a}\|\,\left\|\vec{b}\right\|}\,\|\vec{a}\|\,\frac{1}{\left\|\vec{b}\right\|}=\frac{\left\langle\vec{a},\,\vec{b}\right\rangle}{\left\|\vec{b}\right\|^2}$$

which is equal to $\|\vec{a}\|$ times the cosine of the angle between $\vec{a}$ and $\vec{b}$ times the reciprocal of $\left\|\vec{b}\right\|$ in the case of real-valued $\vec{a}$ and $\vec{b}$ so that, when it is multiplied by $\vec{b}$, it gives $\|\vec{a}\|$ times the cosine of the angle between $\vec{a}$ and $\vec{b}$ times the unit vector in the direction of $\vec{b}$, where $\left\langle\vec{a},\,\vec{b}\right\rangle$ means the Hermitian inner product $\sum_{j=1}^n a_j\overline{b_j}$. This choice of λ can be argued by minimizing the distance $\left\|\vec{a}-\lambda\vec{b}\right\|^2$ or by making $\vec{a}-\lambda\vec{b}$ perpendicular to $\vec{b}$ with respect to the Hermiitan inner product. For that particular value of λ we have the Pythagorean identity

$$\left\|\vec{a}-\lambda\vec{b}\right\|^2=\|\vec{a}\|^2-\left\|\lambda\vec{b}\right\|^2$$

which comes from the expansion

$$\left\|\vec{a}-\lambda\vec{b}\right\|^2=\|\vec{a}\|^2+\left\|\lambda\vec{b}\right\|^2-2\mathrm{Re}\left(\overline{\lambda}\left\langle\vec{a},\,\vec{b}\right\rangle\right)$$

$$= \|\vec{a}\|^2 + \left| \frac{\left\langle \vec{a},\, \vec{b} \right\rangle}{\left\|\vec{b}\right\|^2} \right|^2 \left\|\vec{b}\right\|^2 - 2\mathrm{Re}\left(\overline{\frac{\left\langle \vec{a},\, \vec{b} \right\rangle}{\left\|\vec{b}\right\|^2}} \left\langle \vec{a},\, \vec{b} \right\rangle \right)$$

$$= \|\vec{a}\|^2 - \left| \frac{\left\langle \vec{a},\, \vec{b} \right\rangle}{\left\|\vec{b}\right\|^2} \right|^2 \left\|\vec{b}\right\|^2 = \|\vec{a}\|^2 - \left\|\lambda\vec{b}\right\|^2.$$

The inequality $\left\|\vec{a} - \lambda\vec{b}\right\|^2 \geq 0$ yields

$$\|\vec{a}\|^2 \geq \left| \frac{\left\langle \vec{a},\, \vec{b} \right\rangle}{\left\|\vec{b}\right\|^2} \right|^2 \left\|\vec{b}\right\|^2$$

which is equivalent to

$$\|\vec{a}\|^2 \left\|\vec{b}\right\|^2 \geq \left| \left\langle \vec{a},\, \vec{b} \right\rangle \right|^2,$$

because the case of $\vec{b} = \vec{0}$ can easily be handled separately. Clearly equality holds if and only if $\vec{a} - \lambda\vec{b} = \vec{0}$. This also means that the minimum of $\left\|\vec{a} - \lambda\vec{b}\right\|^2$ is realized by 0 when $\lambda = \frac{\left\langle \vec{a},\, \vec{b} \right\rangle}{\left\|\vec{b}\right\|^2}$.