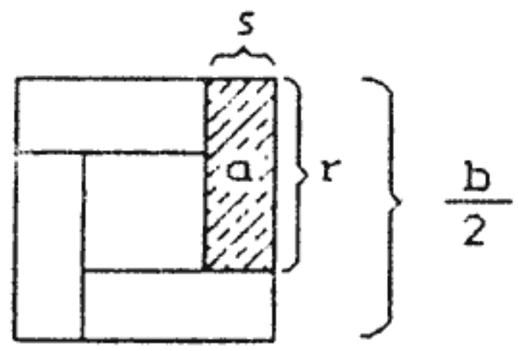
Aspetti didattici della teoria delle equazioni algebriche

Hartmut Laue

Dedicato a Hilger Wolff in occasione del suo 60^{o} compleanno

1. Equazioni quadratiche (motivazione geometrica). Se r, s sono i lati di un rettangolo, allora la sua area è $a = rs$, e il suo perimetro è $b = 2(r+s)$. Consideriamo per il momento la domanda inversa: siano noti l'area a e il perimetro b di un rettangolo; come si determinano i lati ? Ciò può essere ritenuta la versione geometrica della seguente domanda aritmetica che fu risolta già dai babilonesi circa il 1700 a.C.: come determinare due numeri quando siano noti la loro somma e il loro prodotto ? Diamo la motivazione della nostra soluzione in base alla seguente figura:



Il quadrato dato si compone di quattro rettangoli di area a e un quadrato (nel mezzo) con il lato $\frac{b}{2} - 2s$. Quindi vale l'equazione $\left(\frac{b}{2}\right)^2 = 4a + \left(\frac{b}{2} - 2s\right)^2$ la quale è equivalente a

$$\left(\frac{b}{2} - 2s\right)^2 = \left(\frac{b}{2}\right)^2 - 4a \qquad (1)$$

Se b e a sono numeri qualsiasi, allora può ben succedere che la differenza a destra in (1) sia negativa; però, il termine a sinistra, essendo un quadrato, sarà sempre non negativo e quindi,

come conseguenza, dobbiamo notare che non per ogni scelta di a, b (numeri non negativi) esiste un rettangolo con l'area a e il perimetro b. La condizione precisa per l'esistenza di una soluzione è, per (1), che valga la disuguaglianza

$$\left(\frac{b}{2}\right)^2 - 4a \geq 0.$$

(Evidentemente il termine a sinistra è l'area del quadratino nel mezzo del quadrato grande di cui sopra.) Assumiamo quest'ultima, allora esiste un rettangolo come richiesto e possiamo infatti determinare i suoi lati trasformando la (1):

$$(1) \Leftrightarrow \frac{b}{2} - 2s = \sqrt{\left(\frac{b}{2}\right)^2 - 4a} \text{ oppure } \frac{b}{2} - 2s = -\sqrt{\left(\frac{b}{2}\right)^2 - 4a}$$
$$\Leftrightarrow s \in \left\{ \frac{b}{4} + \frac{1}{2}\sqrt{\left(\frac{b}{2}\right)^2 - 4a}\ ,\ \frac{b}{4} - \frac{1}{2}\sqrt{\left(\frac{b}{2}\right)^2 - 4a} \right\}.$$

Dunque, i lati del rettangolo cercato si determinano aggiungendo o togliendo (risp.) a $\frac{b}{4}$ la metà del lato del quadratino nel mezzo della figura. In questo senso, il problema del trovare i lati di un rettangolo viene ridotto a quello del determinare il lato di un quadrato.

Se rinunciamo all'interpretazione dei valori come incognite geometriche e ci concentriamo sulle soluzioni numeriche di (1), allora otteniamo le soluzioni anche nel caso che

$$\left(\frac{b}{2}\right)^2 - 4a < 0:$$

Ricordiamo che, nel campo $\mathbb{C}$ dei numeri complessi, l'equazione $x^2+1=0$ ha due soluzioni di cui una viene denotata con i. Se ammettiamo di ampliare il campo dei numeri reali, accettendo soluzioni in tutto $\mathbb{C}$, allora possiamo andare avanti anche in

questo caso:

$$(1) \Leftrightarrow \frac{b}{2} - 2s = i\sqrt{4a - \left(\frac{b}{2}\right)^2} \text{ oppure } \frac{b}{2} - 2s = -i\sqrt{4a - \left(\frac{b}{2}\right)^2}$$

$$\Leftrightarrow s \in \left\{\frac{b}{4} + \frac{i}{2}\sqrt{4a - \left(\frac{b}{2}\right)^2}\ ,\ \frac{b}{4} - \frac{i}{2}\sqrt{4a - \left(\frac{b}{2}\right)^2}\right\},$$

cioè, in $\mathbb{C}$, il problema aritmetico di cui sopra ha sempre una soluzione e rimane solo la questione se essa sia interpretabile come lato di un rettangolo, cioè: se essa sia reale e positiva. Quest'ultima . domanda ha più il carattere di una richiesta supplementare da soddisfarsi dopo la soluzione aritmetica completa che rappresenta il nocciolo del pensiero, e quindi è ragionevole separare mentalmente le due cose.

I numeri di partenza a $(=rs)$ e b $(=2(r+s))$ appaiono in modo naturale quando si forma il prodotto dei polinomi lineari $x-r$ e $x-s$; evidentemente, il polinomio prodotto

$$f(x) = (x-r)(x-s) = x^2-(r+s)x+rs = x^2-\frac{b}{2}x+a$$

ha precisamente i numeri cercati r, s come zeri[1]. Pertanto possiamo riformulare il nostro problema di partenza così:

Siano dati numeri a, b. Il polinomio $x^2-\frac{b}{2}x+a$ ha zeri ? Se essi esistono, come possiamo descriverli ?

A questo punto possiamo fare un'osservazione importante, un raffinamento, in relazione al nostro modo di domandare: data un'equazione è importante sapere dove cerchiamo le sue soluzioni, se nell'ambito dei numeri razionali, reali, complessi ? Vedremo nel seguito che c'è un gran numero di ambiti numerici rispetto a cui la domanda dell'esistenza di una soluzione può essere

[1]Ricordiamo che un numero a si dice zero del polinomio f se $f(a)=0$

interessante. E ancora un'altro commento: come mai possiamo essere sicuri di trovare almeno un campo di numeri che contiene uno zero del polinomio ? Questa domanda ("assoluta") dell'esistenza di uno zero dovrebbe precedere quella "relativa" dell'esistenza di uno zero in un certo ambito, e questa ultima dovrebbe precedere quella di una certa forma o rappresentazione degli zeri che si manifesta nella caccia di una cosiddetta "formula delle soluzioni". Infatti, fu proprio questa ricerca calcolistica cieca a sbarrare per secoli la strada del successo nell'inseguimento del problema generale della risoluzione delle equazioni. Naturalmente, il caso dell'equazione quadratica era seducente ... ma traeva in inganno! Perciò, anche se le equazioni di grado superiore al secondo non vengono trattate in modo completo nell'insegnamento, oggi è importante evitare nell'impostazione del problema delle equazioni di secondo grado l'ottica ingenua che ha portato fuori strada tutti gli studiosi prima di Galois a partire dai babilonesi.

Possiamo esprimere il nostro risultato fin qui ottenuto nel modo seguente (ponendo $a_1 := -\frac{b}{2}$, $a_2 := a$):

Teorema 1. Siano a_1, a_2 numeri arbitrari. Se $\rho \in \mathbb{C}$ é tale che vale $\rho^2 = a_1^2 - 4a_2$, allora $-\frac{1}{2}a_1+\frac{1}{2}\rho$, $-\frac{1}{2}a_1-\frac{1}{2}\rho$ sono le soluzioni della equazione $x^2+a_1x+a_2 = 0$.
(In particolare, ogni equazione quadratica è risolubile in $\mathbb{C}$.)

In altre parole, per risolvere l'equazione generale di secondo grado $x^2+a_1x+a_2 = 0$ basta risolvere l'equazione quadratica pura (cioè, senza addendo lineare) $x^2-(a_1^2-4a_2) = 0$. Le soluzioni ρ di quest'ultima portano nel modo descritto alle soluzioni dell'equazione generale. Così dobbiamo leggere la "formula delle soluzioni dell'equazione quadratica": essa esprime un teorema di riduzione: il problema dell'equazione generale di secondo grado viene ridotto al caso speciale di una equazione pura di secondo grado. E questa riduzione è la traduzione al livello astratto di

quella geometrica fatta precedentemente "con gli occhi".

2. **Le domande chiave della teoria delle equazioni.** Il nostro problema di partenza ha un analogo spaziale: siano noti il volume a, la superficie b e la somma c degli spigoli di un parallelepipedo. È possibile determinare gli spigoli ?

Come nel caso quadratico, i numeri dati sono, a meno di fattori razionali, i coefficienti di un polinomio cubico i cui zeri sono i tre spigoli r, s, t, perchè vale

$$\begin{aligned}(x-r)(x-s)(x-t) &= x^3 - (r+s+t)x^2 + (rs+rt+st)x - rst \\ &= x^3 - \frac{c}{4}x^2 + \frac{b}{2}x - a.\end{aligned}$$

Se, come sopra, non prendiamo in considerazione per il momento la questione di una possibile interpretazione geometrica dei risultati, il problema si legge così (ponendo $a_1 := -\frac{c}{4}$, $a_2 := \frac{b}{2}$, $a_3 := -a$): il polinomio $x^3+a_1x^2+a_2x+a_3$ ha zeri ? Se essi esistono, come possiamo descriverli ?

Descriviamo nel seguito la soluzione classica del problema, trovata più di 400 anni fa. Poniamo $y := x+\frac{1}{3}a_1$. Allora $x = y - \frac{1}{3}a_1$ e $x^3 + a_1x^2 + a_2x + a_3 = y^3 + b_2y + b_3$, ove $b_2 = a_2 - \frac{1}{3}a_1^2$ e $b_3 = a_3 + \frac{2}{27}a_1^3 - \frac{1}{3}a_1a_2$. Pertanto basta rispondere alla domanda per il polinomio $y^3 + b_2y + b_3$; gli zeri del polinomio originale saranno soltanto additivamente spostati di $\frac{1}{3}a_1$. Poi, il problema principale è trovare uno zero qualunque; trovato un tale σ, possiamo dividere il polinomio dato per $y-\sigma$, e gli ulteriori zeri saranno gli zeri del polinomio quadratico che risulta, cioè, saranno determinabili col metodo trattato precedentemente. Benché i babilonesi sapessero risolvere le equazioni di secondo grado nel 1700 a. C., non prima del sedicesimo secolo gli algebristi riuscirono a risolvere le equazioni cubiche in generale. La soluzione, trovata da Tartaglia, inoltrata a Cardano e pubblicata da esso sul libro "Ars Magna" nel 1545, è esprimibile in questa forma:

Teorema 2 (Scipione del Ferro ?, Tartaglia; Cardano 1545). Siano b_2, b_3 arbitrari numeri, $b_2 \neq 0$. Sia δ una soluzione della equazione quadratica pura

$$x^2-(\frac{1}{27}b_2^3+\frac{1}{4}b_3^2) = 0.$$

Sia ζ una soluzione dell'equazione cubica pura

$$x^3+(\frac{1}{2}b_3-\delta) = 0.$$

Allora vale $\zeta \neq 0$, e $\zeta-\frac{1}{3\zeta}b_2$ è una soluzione dell'equazione

$$y^3+b_2y+b_3 = 0.$$

(In particolare, ogni equazione cubica è risolubile in $\mathbb{C}$.)[2]

Anche questo risultato è una riduzione alla soluzione di equazioni pure: si ottiene una soluzione della equazione cubica generale risolvendo prima un'equazione quadratica pura e successivamente un'equazione cubica pura (cioè, un'equazione cubica in cui sono nulli i coefficienti della prima e della seconda potenza dell'indeterminata). Infatti, la soluzione δ della equazione quadratica appare nel coefficiente dell'equazione cubica pura da risolvere dopo.

A questo punto le nostre considerazioni danno luogo alle seguenti domande chiave della teoria delle equazioni:

[2]Evitiamo di utilizzare il simbolo della radice ($\sqrt{\ }$) per i numeri complessi perché non c'è nessuna possibilità ragionevole di associare un unico numero complesso ad esso. Quasi tutti i testi che trattano le equazioni cubiche utilizzando quel simbolo si astengono del discorso spinoso ma sostanziale della scelta giusta della radice, lasciando spazio a interpretazioni dei termini indicati che non portano a una soluzione dell'equazione da risolvere.

Sia $n \in \mathbb{N}$ e siano $a_1,\ldots,a_n$ numeri qualsiasi.

1. L'equazione

$$(*) \qquad x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n = 0$$

è risolubile in $\mathbb{C}$?

2. Se la risposta di **1.** è positiva, è possibile ridurre la soluzione di (*) alla determinazione delle soluzioni di opportune equazioni pure di grado k con $2 \leq k \leq n$?

Sottolineamo che queste domande sono molto più prudenti di quella che richiede una "formula" delle soluzioni, domanda che sottintenderebbe senz'altro l'esistenza di una tale espressione esplicita. Proprio questa sarà un grosso problema! E anche se avessimo già superato quest'ostacolo, sarebbe ancora lontana una formula come quella che abbiamo già a disposizione (Teoremi 1 e 2) nei casi delle equazioni quadratiche e cubiche per le quali, infatti, la soluzione generale si compone di "radici" (cioè, di zeri di polinomi "puri" del tipo x^k+c).

Poco dopo la scoperta della soluzione delle equazioni cubiche un allievo di Cardano, Ferrari, riuscì a risolvere le equazioni di quarto grado, riducendole alle equazioni di terzo grado. Per $n \geq 5$ invece le domande **1.**, **2.** rimasero aperte ancora più di 250 anni.

Durante una lunga epoca il problema delle equazioni fu considerato sotto un'ottica che generalizza la nostra introduttiva formulazione geometrica della domanda: cioè, assumendo uno spezzamento completo di un polinomio in fattori lineari, i suoi coefficienti sono costituiti in modo molto regolare dai suoi zeri, il che viene espresso nel seguente teorema già noto a Viète (1540-1603) e pubblicato dopo la sua morte:

Teorema 3 (il cosiddetto teorema di Viète, 1615). Sia $n \in \mathbb{N}$ e siano $a_1,\ldots,a_n$ numeri qualsiasi. Se $\alpha_1,\ldots,\alpha_n$ sono numeri tali che vale

$$(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n) = x^n+a_1x^{n-1}+a_2x^{n-2}+\cdots+a_{n-1}x+a_n,$$

allora si ha

$$\begin{aligned}
\alpha_1 + \alpha_2 + \cdots + \alpha_n &= -a_1\\
\alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_{n-1}\alpha_n &= a_2\\
\alpha_1\alpha_2\alpha_3 + \alpha_1\alpha_2\alpha_4 + \cdots + \alpha_{n-2}\alpha_{n-1}\alpha_n &= -a_3\\
&\vdots\\
\alpha_1\cdots\alpha_n &= (-1)^n a_n.
\end{aligned}$$

Dunque i coefficienti di un polinomio comportano un'interessante informazione sugli zeri in questione: a meno del segno, il coefficiente a_j è la somma dei prodotti di lunghezza j formati dagli zeri cercati $\alpha_1,\ldots,\alpha_n$. Questa interpretazione dei coefficienti del polinomio comporta anche una spiegazione perchè, sopratutto nella bibliografia vecchia, gli indici dei coefficienti crescono quando gli esponenti di x decrescono: così vengono associati i prodotti di j zeri alla potenza x^j. Indubbiamente oggi è passata in seconda linea questa interpretazione dei coefficienti di un polinomio nello spirito del teorema di Viète che una volta fu considerato come fondamentale. I testi moderni comunque preferiscono, per molti motivi pratici, la scrittura $x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$.

Dopo secoli di sforzo, finalmente furono trovate le risposte definitive delle domande **1.** e **2.** anche per $n \geq 5$. Bisognò aspettare nientemeno che C. F. Gauss il quale dimostrò per primo nel 1799 che la domanda **1.** ha una risposta positiva; all'epoca il problema risolto fu considerato tanto centrale e il risultato tanto significativo che quest'ultimo viene chiamato, per motivi storici anche nei nostri giorni, "il teorema fondamentale dell'Algebra". Infatti si tratta di uno dei risultati più belli e soddisfacenti della matematica:

Teorema 4 (il cosiddetto teorema fondamentale dell'Algebra, Gauss 1799). Siano $n \in \mathbb{N}$ e siano $a_1,\ldots,a_n \in \mathbb{C}$. Allora l'equazione $x^n+a_1x^{n-1}+a_2x^{n-2}+\cdots+a_{n-1}x+a_n = 0$ è risolubile in $\mathbb{C}$.

Contrariamente alle speranze plurisecolari, la risposta alla domanda **2.** è negativa per $n \geq 5$. Completando gli studi di P. Ruffini, il norvegese Niels Henrik Abel (1802-1829) dimostrò che non è possibile determinare le soluzioni di un'equazione generale data risolvendo consecutivamente equazioni pure (cioè, estraendo radici). Cioè, dal famoso teorema di Abel e Ruffini (1824) consegue che, nel caso generale, non esiste nessuna strada per esprimere gli zeri come composizione idonea di radicali: una formula come nei casi $n = 2, 3, 4$ non esiste per alcun $n \geq 5$. Un esempio di un polinomio i cui zeri non sono riducibili a radicali è x^5-4x+2; torneremo in seguito su ciò.

Tutto sommato, lo sviluppo della teoria delle equazioni fino al teorema di Abel-Ruffini viene schematizzato con la seguente tabella:

n	1.	2.			
2	sì	sì	1700 a.C.	babilonesi	Teorema 1
3	sì	sì	<1526	Scipione del Ferro (Bologna), non pubbl.	
			1545	Tartaglia (Venezia)	Teorema 2
				Cardano (Milano)	"Ars Magna"
4	sì	sì	1545	Ferrari	"Ars Magna"
≥ 5	sì		1799	Gauss	Teorema 4 (il teorema fondam. dell'Algebra)
		no	1824	Abel-Ruffini	

Il teorema di Abel-Ruffini pose fine alla ricerca di una formula delle soluzioni analoga alle formule per i gradi 2, 3, 4. Però, ci sono naturalmente moltissime equazioni di grado superiore i cui zeri sono determinabili estraendo idonee radici una dopo l'altra. Tali equazioni si dicono "risolubili per radicali". Questa nozione sarà precisata ben presto, essendo il fulcro di

quanto segue. Siamo arrivati al punto in cui l'oggetto principale della teoria di Galois prende profilo. Superando notevolmente la domanda **2.**, la teoria di Galois prende la risposta negativa "globale" del teorema di Abel-Ruffini come motivazione della domanda "locale":

3. Quando un'equazione è risolubile per radicali ?

Se, secondo il teorema di Abel-Ruffini, non c'è nessuna formula generale delle soluzioni per $n \geq 5$, sarà almeno possibile caratterizzare i polinomi che sono risolubili per radicali ?

3. Precisazione della domanda fondamentale della teoria di Galois.

La domanda è quando gli zeri di un polinomio sono rappresentabili formando somme, differenze, prodotti, quozienti e radicali (anche ripetutamente), partendo dai suoi coefficienti. Le prime quattro operazioni sono quelle in un campo, mentre i radicali sono le soluzioni delle equazioni pure, cioè delle equazioni del tipo

$$x^k + a = 0.$$

Se K è un campo, $\mathbb{Q} \leq K \leq \mathbb{C}$, $a \in K$ e ρ è una soluzione di questa equazione, allora gli elementi

$$b_0 + b_1\rho + \cdots + b_{k-1}\rho^{k-1} \quad (b_0,\ldots,b_{k-1} \in K)$$

formano un sottocampo K' di $\mathbb{C}$ che contiene K. Un tale campo si dice una estensione di K mediante un radicale, perché esso è il minimo sottocampo di $\mathbb{C}$ che contiene K e quel radicale (ρ). I numeri che si ottengono estrando successivamente radici (e collegandole con i numeri già ottenuti precedentemente mediante le leggi di operazione di campo) sono evidentemente tutti e soli quelli che nascono come elementi di estensioni mediante radicali ripetute.

Definizione. Un sovracampo L di $\mathbb{Q}$ si dice una estensione mediante radicali successivi di $\mathbb{Q}$ se esiste una catena finita

$$\mathbb{Q} = K^{(0)} \subseteq K^{(1)} \subseteq \cdots \subseteq K^{(m)} = L$$

di campi $K^{(j)}$ tali che $K^{(j)}$ è un'estensione mediante un radicale di $K^{(j-1)}$, per $1 \leq j \leq m$.

Nel seguito ci limitiamo al caso classico di un polinomio su $\mathbb{Q}$ e quindi precisiamo la domanda 3. in questo modo:

4. Sia $f(x)$ un polinomio a coefficienti razionali. Quando esiste un'estensione mediante radicali successivi L di $\mathbb{Q}$ tale che, per elementi idonei $\alpha_1,\ldots,\alpha_n \in L$ (che non sono necessariamente distinti), vale $f(x) = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$?

Per il Teorema 4 esistono sicuramente elementi $\alpha_1,\ldots,\alpha_n \in \mathbb{C}$ tali che $f(x) = (x-\alpha_1)\cdots(x-\alpha_n)$. La domanda, però, se questi zeri $\alpha_1,\ldots,\alpha_n$ siano esprimibili utilizzando solo le operazioni nel campo e radici consecutivamente estratte è, in forma precisa, esattamente la domanda se gli zeri appartengano a una estensione mediante radicali successivi di $\mathbb{Q}$. Né la domanda dell'esistenza degli zeri né quella del loro calcolo numerico è il punto saliente della teoria di Galois, ma lo è la questione della loro eventuale esprimibilità nel modo descritto all'inizio di questo capitolo.

A questo punto sarà opportuno chiarire preventivamente un possibile equivoco: senz'altro esiste un sottocampo minimo di $\mathbb{C}$ che contiene tutti gli zeri $\alpha_1,\ldots,\alpha_n$, esso è l'intersezione di tutti i sottocampi che li contengono; denotiamolo nel seguito con $\mathbb{Q}(f)$ e chiamiamolo il campo di spezzamento di f in $\mathbb{C}$. Se l'equazione data è risolubile per radicali, allora $\mathbb{Q}(f)$ deve essere contenuto in un'estensione mediante radicali successivi L di $\mathbb{Q}$. È importante notare, però, che questo può essere il caso senza che $\mathbb{Q}(f)$ stessa

sia un'estensione mediante radicali successivi. Quindi non è sufficiente controllare se quel campo minimo $\mathbb{Q}(f)$ stesso sia una estensione mediante radicali successivi di $\mathbb{Q}$. In sostanza, questo era già noto a Cardano. Alcuni polinomi cubici irriducibili hanno tre zeri reali; quindi, il loro campo di spezzamento è contenuto nel campo $\mathbb{R}$. Tuttavia si dimostra che una rappresentazione degli zeri con l'aiuto di radicali è soltanto possibile utilizzando numeri immaginari: nonostante il fatto che esista un'estensione mediante radicali successivi L di $\mathbb{Q}$ contenente tutti gli zeri, il campo di spezzamento non ha questa proprietà, esso è troppo piccolo. Questi casi, in cui gli zeri, pur essendo reali, sono rappresentabili mediante radicali soltanto con l'aiuto di numeri immaginari, erano accettati solo con sospetto all'epoca di Cardano, perché ci si fidava di zeri reali ma non di numeri immaginari. Il caso descritto di un'equazione cubica con tre zeri reali si dice il "casus irreducibilis"; un tale esempio è l'equazione $x^3-3x-1=0$. Risolviamola applicando il teorema 2: poniamo $\delta := \frac{i}{2}\sqrt{3}$ e scegliamo un numero complesso ζ tale che $\zeta^3 = \delta+\frac{1}{2}$. Allora, per il teorema 2, il numero $\zeta-\frac{1}{3\zeta}\cdot(-3)$ è una soluzione dell'equazione, ed è uguale a $\zeta+\frac{1}{\zeta}$. È facile calcolare che vale $\zeta^{18} = 1$, e quindi ζ è una radice diciottesima dell'unità. In particolare vale $\frac{1}{\zeta} = \bar{\zeta}$ (il coniugato di ζ) e quindi la soluzione trovata, essendo uguale a $\zeta+\bar{\zeta}$, è un numero reale, scritto come somma di due radicali non reali. Questo fenomeno è tipico per il casus irreducibilis: la soluzione dell'equazione cubica in questione è la somma di due radicali complessi le cui parti immaginarie sono opposte e quindi si semplificano. Pur essendo interessati solo alle parti reali abbiamo nondimeno bisogno di numeri non reali per rappresentare le soluzioni tramite radicali; a tale scopo, né gli elementi del campo di spezzamento e nemmeno quelli di $\mathbb{R}$ bastano.

Pertanto rendendoci conto del fatto che non basta chiedere se il campo di spezzamento di un polinomio sia un'estensione mediante radicali successivi, possiamo comunque riformulare **4.** così:

5. Sia f un polinomio a coefficienti razionali. Quando esiste una estensione mediante radicali successivi L di $\mathbb{Q}$ che contenga il campo di spezzamento $\mathbb{Q}(f)$?

4. **La risposta di Galois.** Sebbene abbiamo visto che bisogna essere prudenti nella formulazione di **5.** e parlare di un campo L che possibilmente è più grande di $\mathbb{Q}(f)$, Evariste Galois (1811-1832) mostrò che, per dare una risposta a **5.**, basta studiare $\mathbb{Q}(f)$ stesso. La sua grande scoperta è un criterio necessario e sufficiente come richiesto in **5.** che discende dallo studio degli automorfismi del campo $\mathbb{Q}(f)$. Gli automorfismi del campo $\mathbb{Q}(f)$ (cioè, gli isomorfismi di $\mathbb{Q}(f)$ su se stesso) formano un gruppo, detto il gruppo di Galois del polinomio f (su $\mathbb{Q}$). Nel 1829, Abel pubblicò un risultato sulla risolubilità di una equazione nel quale l'ipotesi sostanziale è la permutabilità di certe trasformazioni del campo di spezzamento. Il suo risultato è contenuto nel seguente

Caso speciale del teorema di Galois. Se il gruppo di Galois di f è commutativo, allora l'equazione $f(x)=0$ è risolubile per radicali.

In riconoscimento della scoperta di Abel che la permutabilità di certi operatori su $\mathbb{Q}(f)$ è una condizione di rilievo nell'ambito del problema, fino a oggi è una usanza comune utilizzare il termine "gruppo abeliano" come sinonimo di "gruppo commutativo". Indipendentemente dagli studi di Abel, Galois dette la risposta completa alla domanda **5.**:

Teorema 5 (Galois 1829/1831). L'equazione $f(x) = 0$ è risolubile per radicali se e solo se nel gruppo di Galois G di f esiste una catena

$$1 = N_0 \leq N_1 \leq N_2 \leq \cdots \leq N_l = G$$

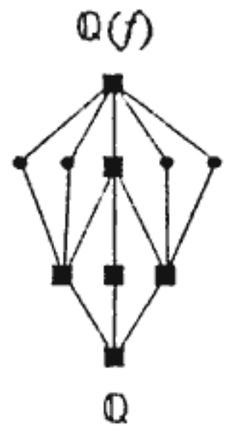
di sottogruppi normali N_j di G tali che i gruppi quoziente N_1/N_0, N_2/N_1, ..., N_l/N_{l-1} sono tutti abeliani.

A causa di questo teorema principale, nella teoria dei gruppi moderna si dice risolubile un gruppo che soddisfa alla proprietà appena descritta in esso. Se il gruppo G è abeliano, allora già la catena corta $1=N_0 \leq N_1=G$ (cioè, $l=1$) ha la proprietà richiesta, e quindi G è risolubile. La risolubilità è una nozione molto più generale della commutatività; gruppi risolubili sono "composti" da gruppi abeliani come descritto nella condizione data nella formulazione del teorema 5. Così si vede che l'idea di Abel introduce in un certo senso alla situazione di base in cui si ha risolubilità per radicali quella in cui G è commutativo. I gruppi risolubili invece nascono "sovrapponendo" i gruppi abeliani (nella forma di una catena), e quindi la nozione di risolubilità si sviluppa dalla nozione di commutatività. Passando dalla commutatività alla risolubilità la condizione sufficiente del suddetto caso speciale si trasforma in una condizione sufficiente e necessaria, quindi in una vera caratterizzazione. Il fatto che la risolubilità del gruppo di Galois è necessaria per poter risolvere l'equazione per radicali rende possibile dimostrare che gli zeri di certi polinomi (di grado ≥ 5 naturalmente) non sono rappresentabili per mezzo di radicali. Infatti, il gruppo di Galois del polinomio $f(x) = x^5-4x+2$ già menzionato sopra è isomorfo al gruppo delle permutazioni di 5 cifre (il "gruppo simmetrico" S_5) per il quale non è molto difficile vedere che non è risolubile. Pertanto siamo sicuri che, grazie al teorema di Galois, l'equazione $x^5-4x+2 = 0$ non è risolubile per radicali.

Il centro della teoria di Galois che rende possibile la suddetta soluzione completa del problema delle equazioni è il cosiddetto teorema principale della teoria di Galois. Una versione debole di esso che, però, permette di farsi un'idea del carattere del risultato dice: Siano f un polinomio a coefficienti razionali, $\mathbb{Q}(f)$ il suo campo di spezzamento in $\mathbb{C}$, G il gruppo di Galois

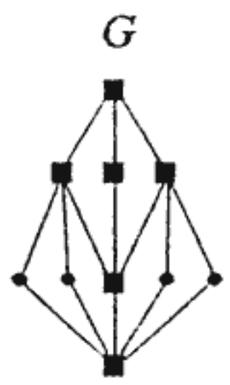
relativo. Allora il reticolo dei sottocampi di $\mathbb{Q}(f)$ è antisomorfo al reticolo dei sottogruppi di G. I sottocampi di $\mathbb{Q}(f)$ che per conto loro sono campi di spezzamento (di altri polinomi) corrispondono ai sottogruppi normali mediante quell'antisomorfismo.

Vogliamo illustrare con un esempio (ma senza dimostrazioni) il contenuto di questo grande teorema. Sia $f(x) := x^4-3$. Il seguente grafico descrive il reticolo dei sottocampi di $\mathbb{Q}(f)$:



(I nodi rappresentano sottocampi di $\mathbb{Q}(f)$; i sottocampi che sono campi di spezzamento sono indicati con un quadratino nero (■), gli altri invece con un punto (•). I segmenti ascendenti indicano l'inclusione.)

Il teorema dice che si ottiene il diagramma per il reticolo dei sottogruppi del gruppo di Galois G di f <u>capovolgendo il diagramma dei sottocampi di</u> $\mathbb{Q}(f)$:



(I punti (•) rappresentano sottogruppi non normali, i quadratini (■) sottogruppi normali; segmenti ascendenti indicano l'inclusione.- In questo caso, il gruppo di Galois G è un cosiddetto "gruppo diedrale" di 8 elementi.)

La formulazione completa del teorema principale della teoria di Galois contiene una descrizione dettagliata della corrispondenza indicata fra i sottocampi di $\mathbb{Q}(f)$ e i sottogruppi

di *G*. Per le dimostrazioni e per ulteriori precisazioni di questa illustre teoria dobbiamo rimandare alle impostazioni in adatti libri di testo (per esempio, [B] o, per dare un libro nella lingua italiana, [GI]). Un'analisi dei contenuti matematici relativi sotto l'aspetto storico (ma non aneddotico) si trova su [E]. Poi, in modo conciso, [W] mette mano a un inquadramento della teoria delle equazioni nello sviluppo grande dell'Algebra.

5. **Conseguenze sotto il punto di vista didattico.** Saltano all'occhio almeno tre punti di contatto della teoria delle equazioni con i contenuti dell'insegnamento a scuola, ciò sono i capitoli che trattano

(i) l'equazione quadratica

(ii) l'introduzione dei numeri irrazionali (il passo al di là di $\mathbb{Q}$)

(iii) il numero π (e il numero e), un numero trascendente

È vero che la formula delle soluzioni di un'equazione quadratica rappresenta uno strumento al cento per cento fidato per trovare gli zeri di un qualunque polinomio di grado 2 a coefficienti numerici. Però, se il nostro scopo fosse quello di determinare numericamente i numeri decimali che risolvono un'equazione data, allora avremmo a disposizione altri metodi più efficaci in quanto che essi funzionerebbero anche per i polinomi di grado superiore: i metodi di approssimazione. Quindi, un tale scopo non può giustificare a fondo il trattamento della formula delle soluzioni di un'equazione quadratica; tuttavia osserviamo naturalmente che essa consente di trovare tali risultati approssimati: chi non vuole (o non sa?) estrarre la radice che compare in essa e fare i calcoli restanti a mano può senz'altro scegliere la strada banale, fidandosi della sua calcolatrice, e arriverà al numero decimale desiderato; anche se così il suo risultato non sarà (quasi) mai

esatto, la precisione sarà sufficiente per la maggior parte delle applicazioni pratiche. Non è "vietato" sfruttare la formula in questa maniera, ma è evidente che il suo valore didattico non si esaurisce con questo.

Nonostante il fatto che quei numeri decimali non sono esatti molti accettano stranamente soltanto essi come risultato e considerano il termine dato dalla formula stessa (che è veramente esatto !) come un numero ancora "da calcolare". Avendo il termine che discende dalla formula, è sempre possibile derivare un'approssimazione decimale, ma viceversa nessun'approssimazione permette di risalire alla rappresentazione esatta (che coinvolge la radice "non ancora calcolata" nel senso suddetto). Ma, alla luce della teoria precedentemente schizzata, non si tratta soltanto di una perdita di precisione. Passando, per esempio, da $\frac{1}{2}(-5+\sqrt{53})$ a 1,14, abbiamo perso molto di più: non ci si accorge più del fatto che il nostro numero è una composizione di numeri razionali con la radice di 53, anzi, tutte le tracce della sua vera natura sono sparite. Ingannati dall'unico scopo di situare il numero sulla retta dei numeri reali, abbiamo sacrificato il carattere individuale del nostro numero a un idolo grigio - alla forma decimale, nella quale, comunque se il numero delle cifre viene limitato, i numeri reali, privati delle loro caratteristiche, vengono costretti ad un'inespressiva anonimità. Il nostro numero invece vive della sua interpretazione come soluzione dell'equazione $x^2+5x-7 = 0$; la radice nel suo interno appare geometricamente come lato del quadrato di cui abbiamo parlato durante la sezione introduttiva di questo articolo. Il numero non è razionale ma coinvolge soltanto una radice quadrata, quindi la sua "altezza" sopra $\mathbb{Q}$ è bassa; infatti, si parla di un elemento "di grado 2 su $\mathbb{Q}$". In questo senso, il numero è "più vicino a $\mathbb{Q}$" che, per esempio, una radice cubica o una radice di un grado ancora più alto. Così si apre un ampio campo di studio: quello dei numeri irrazionali. Tutto il sottile tessuto del mondo dei

numeri che traspare a questo punto viene strappato di colpo se si mette troppo in evidenza la forma decimale. Lo sviluppo della formula delle soluzioni dell'equazione quadratica invece dà un'ottima occasione per portare la direzione dello sguardo, con molta cautela, verso quel mondo affascinante. Per gli studenti si tratta di un passo significativo fuori di $\mathbb{Q}$; è importante farlo con circospezione: la motivazione, per il momento, è quella del risolvere un'equazione ma, senza che lo studente lo presagisse, il risultato comporta un accesso a una sottoclasse dei complessi di caratteristica molto particolare. Il vero compito didattico quindi non è esercitare l'automatismo dall'equazione data fino alla soluzione in forma decimale, ma tirare alla luce il fatto che queste soluzioni sono, come numeri, di una natura molto speciale che merita essere investigata in dettaglio, distinguendoli in modo chiaro dall'insieme molto più vasto di tutti i numeri complessi (o reali); questo è possibile, per esempio mediante una interpretazione geometrica o, più in termini algebrici, mettendo in evidenza l'"altezza" della radice (=2) e vuole senz'altro il contrasto con i numeri che non sono di questa forma speciale.

Conviene precisare a questo punto cosa intendiamo parlando del "carattere" e della "natura" di un numero. Dopo $\mathbb{N}$, $\mathbb{Z}$ e $\mathbb{Q}$ incominciano i problemi sottili relativi alla nozione di numero; confrontato con il passaggio da $\mathbb{N}$ a $\mathbb{Z}$ e da $\mathbb{Z}$ a $\mathbb{Q}$, quello da $\mathbb{Q}$ a $\mathbb{R}$ (o a $\mathbb{C}$) è sproporzionatamente maggiore, sia nell'ottica storica che sistematica che didattica. La teoria delle equazioni comporta un certo orientamento nell'ambito vasto dei numeri irrazionali: essa si occupa dei numeri che sono zeri di polinomi non nulli; tali numeri si dicono "numeri algebrici" e formano un sottocampo $\mathbb{A}$ di $\mathbb{C}$. Denotando con $\mathbb{W}$ [3] l'insieme dei numeri che appartengono a una estensione mediante radicali successivi, possiamo esprimere la domanda, aperta fino ad Abel, così: vale $\mathbb{W} = \mathbb{A}$? Si ha $\mathbb{W} \subseteq \mathbb{A}$, ma

[3] dal tedesco "Wurzel" (radicale)

questa inclusione è propria: vale $\mathbb{W} \subset \mathbb{A}$; ci sono numeri algebrici che non appartengono a $\mathbb{W}$, vedi il "risultato negativo" di Abel. Il teorema 5 di Galois invece da, in termini gruppali, una caratterizzazione di $\mathbb{W}$ come sottoinsieme di $\mathbb{A}$. La conseguenza storica dei risultati di Galois fu la nascita di un'altra grande teoria algebrica: la teoria algebrica dei numeri; il cui oggetto classico, detto in modo molto accorciato, è lo studio del campo $\mathbb{A}$.

Ma il passo da $\mathbb{Q}$ via $\mathbb{W}$ ad $\mathbb{A}$, pur comprendendo una ricchezza incredibile di fenomeni che meritano uno studio profondo, è ancora piccolo per quanto riguarda la cardinalità degli insiemi considerati: gli insiemi $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{W}$, $\mathbb{A}$ sono tutti numerabili. Cioè, sotto l'ottica delle cardinalità, la parte del leone dello insieme dei numeri complessi (e anche dei reali) è l'insieme dei numeri trascendenti (cioè, non algebrici) che non abbiamo nemmeno toccato malgrado l'ampiezza di tutti gli argomenti già sfiorati fin qui. Infatti è molto più facile dimostrare che l'insieme dei numeri trascendenti non è numerabile che dimostrare in un solo caso che un certo numero concreto non sia algebrico. Ci sono numeri trascendenti in eccesso - ma è difficilissimo individuarne uno. Sono passati solo poco più di cento anni dalle prime dimostrazioni della trascendenza dei numeri e (Hermite, 1873) e π (Lindemann, 1882).

Lo studio dei numeri irrazionali, evidentemente, non può essere il compito dell'insegnamento scolastico: comunque in senso positivo non sarà possibile introdurre gli insiemi $\mathbb{W}$, $\mathbb{A}$ e parlare in modo consolidato della trascendenza. Però, in senso negativo, consci dei risultati della teoria delle equazioni, si possono combattere possibili idee sbagliate sui numeri, prevedendo, per evitarli, gli eventuali corti circuiti mentali degli studenti, fra i quali i più diffusi sono i seguenti:

considerare uguali i concetti di "numero irrazionale" e "radice" (o, ancora peggio, "radice quadrata")

considerare uguali i concetti di "numero irrazionale" e "zero (non razionale) di un polinomio non nullo" (cioè, "numero algebrico non razionale")

accettare soltanto la forma decimale per rappresentare un numero reale (o, ancora peggio, la forma decimale con un numero finito di cifre, oppure, pessimo, la forma decimale con un numero fisso di cifre).

Per esempio, è importante parlare non solo delle radici quadrate ma (almeno) anche delle radici cubiche per far vedere che esse sono numeri "nuovi". Non è esclusa una dimostrazione che $\sqrt[3]{5}$ non coincide con nessuna radice quadrata di un numero razionale; il fatto che essa non è nemmeno rappresentabile come combinazione di radici quadrate a coefficienti razionali sarà troppo difficile ... ma non dovrebbe mancare come informazione! Questa "strategia" è sempre consigliabile quando l'argomento è veramente profondo, troppo profondo per un trattamento comprensibile da parte degli studenti: studiare in dettaglio un pezzettino intellettualmente raggiungibile del fenomeno generale in modo che si apra lo sguardo; dopo tale apertura dell'orizzonte il tempo può essere maturo (la decisione del momento giusto dipende naturalmente dalle reazioni degli studenti) per dare un risultato semplicemente come informazione. Questo procedere non deve essere confuso con un cattivo "insegnamento cattedratico": il centro, naturalmente, è sempre la sensibilizzazione degli studenti per l'argomento, e solo dopo ciò l'insegnante tiene conto della curiosità risvegliata, dando un risultato che, ragionando con onestà didattica, risulterebbe ancora fuori portata. Come sempre, la nascita della domanda è lo scopo principale, non il memorizzare delle risposte. Così è anche possibile discutere la rappresentazione cartesiana di un polinomio di quinto grado come x^5-4x+2 per evidenziare in modo

grafico dove essa intersechi l'asse orizzontale, anche per tentare un'approssimazione dello zero relativo. Dopo sarà una grande sorpresa, una delusione matematicamente giustificata, venire a sapere che non c'è nessuna possibilità di calcolare essa tramite radici. Gli studenti capiranno benissimo questo modo di dire un po' impreciso, con il riferimento alla soluzione dell'equazione quadratica. Gli occhi "vedono" lo zero del polinomio, ma si tratta di un numero fuori del mondo delle radici: "le radici", quindi, sono numeri particolari da non confondere con la totalità dei numeri... Finalmente, tutto il trattamento dettagliato del numero π dovrebbe lasciare spazio anche per mettere in evidenza che si tratta di un numero che non risulta mai come soluzione di una equazione algebrica: dopo tanti "acquisti" di nuovi numeri reali mediante le soluzioni di equazioni stiamo studiando un numero di massima importanza per la geometria elementare e per la trigonometria la cui natura è completamente diversa: lontana dalle radici, lontana (più in generale) dagli zeri dei polinomi...

Quando la materia da insegnare è esigente come in questo caso, sarebbe sbagliata una decisione didattica che dia come risultato un'idea storpiata dei concetti. Il principio sano che l'insegnamento di matematica non sopporta il semplice dare ("dettare") risultati senza un ragionamento adeguato d'altronde non può giustificare che, dopo tanti anni a scuola, il concetto di numero venga confuso nel senso suddetto. La necessità di trattare $\mathbb{R}$ (o $\mathbb{C}$) comporta, purtroppo, che dobbiamo trasgredire quel principio, come sempre quando contemporaneamente un punto è matematicamente troppo profondo per uno svolgimento soddisfacente e, sopprimendolo, metteremmo in pericolo il tutto. Abbiamo, però, anche tentato di indicare come una tale trasgressione può succedere con responsabilità.

Le conseguenze didattiche della teoria di Galois (per l'insegnamento alla scuola media) riguardano quindi uno dei due concetti originari della matematica: quello di NUMERO (l'altro è

senza dubbio quello di SPAZIO). Dovrebbe essere uno dei più nobili compiti dell'insegnamento della matematica sviluppare una viva coscienza delle proprietà e caratteristiche dei numeri, visto che la loro natura e i loro misteri hanno spinto e diretto gli sforzi dei matematici di ogni epoca. Lo studio della teoria delle equazioni, con la teoria di Galois come culmine, fornisce un ottimo orientamento su quell'argomento fondamentale. Dovrà essere diversa da caso a caso la decisione fino a quale punto, con quale misura di profondità e sottigliezza, sia ragionevole prefiggersi degli scopi didattici nello spirito di questo articolo, in dipendenza dagli ulteriori fattori pedagogici generali del momento. Auguriamoci che tale decisione sia sempre sorretta dalla convinzione che le nozioni devono precedere e giustificare gli algoritmi, che il pensare deve dominare e determinare il fare, che la autonomia mentale come scopo pedagogico è di gran lunga superiore all'apprendimento delle capacità esecutive, anche se perfette.

Bibliografia

[B] J. R. Bastida, Field extensions and Galois theory, Addison-Wesley Publ. Comp., Reading (Mass.) 1984

[E] H. M. Edwards, Galois theory, Springer, New York - Berlin - Heidelberg 1984

[GI] M. Girardi, G. Israel, Teoria dei campi, Feltrinelli edit., Milano 1976

[W] B. L. van der Waerden, A history of algebra, Springer, Berlin - Heidelberg 1985.