Dimostrazione del Teorema di Natale di Fermat

Serie sul Teorema di Natale di Fermat

2025-12-25

ITA ENG

Teorema di Natale

Ricapitoliamo ciò che abbiamo visto finora. Iniziamo enunciando il teorema di Natale in termini di congruenze.

Theorem 0.1 (Teorema di Natale)**.** *Un numero primo dispari p si può scrivere come $p = a^2 + b^2$ con $a, b \in \mathbb{Z}$ se e solo se $p \equiv_4 1$.*

Nel post precedente abbiamo dimostrato che se p si può scrivere come somma di due quadrati, allora ha resto 1 nella divisione per 4, cioè $p \equiv_4 1$. Ora vogliamo dimostrare il viceversa: se $p \equiv_4 1$, allora p si può scrivere come somma di due quadrati (questa è la parte difficile).

In questo post sviluppiamo tutto ciò che serve per dare una dimostrazione completa del teorema. Per prima cosa spostiamo il problema nel contesto degli interi di Gauss, che sono un'estensione degli interi che già conosciamo. Poi vedremo come le somme di due quadrati compaiono in modo naturale e come riconoscere i primi che si possono scrivere come somma di due quadrati.

Interi di Gauss

Sappiamo tutti che l'equazione

$$x^2 + 1 = 0 \tag{1}$$

non ha soluzioni negli interi, perché i quadrati sono sempre non negativi e, aggiungendo 1 a un numero non negativo, non si può ottenere zero.

Aggiungiamo allora una soluzione di questa equazione agli interi: definiamo

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

così che

$$i^2 + 1 = \sqrt{-1}^2 + 1 = -1 + 1 = 0.$$

Quindi ora abbiamo una soluzione di Equation 1, che è i.

Definiamo $\mathbb{Z}[i]$ (gli interi di Gauss) come gli interi insieme a questo i. I "numeri" in $\mathbb{Z}[i]$ sono della forma

$$a + b \cdot i$$

dove a e b sono interi usuali. Per esempio,

$$1 + 2i, \qquad 5 + 7i, \qquad 0 + i$$

sono tutti interi di Gauss. I casi particolari della forma $a + 0i$ li denoteremo semplicemente con a, mentre i numeri $0 + bi$ li denoteremo con bi. In particolare, osserviamo che dentro $\mathbb{Z}[i]$ ci sono tutti gli interi: infatti li recuperiamo come numeri della forma $a + 0i = a$.

Ma perché dovremmo lavorare negli interi di Gauss invece che negli interi "classici"? Cercheremo di capirlo lungo la dimostrazione del Teorema di Natale. Il punto chiave è che le soluzioni dell'equazione polinomiale

$$x^2 + 1 = 0$$

sono strettamente legate al nostro teorema.

Ma prima dobbiamo capire come funzionano somme e prodotti negli interi di Gauss.

È semplice: basta trattare i come una variabile formale e ricordare che $i^2 = -1$. Spiego meglio:

Supponiamo di voler sommare $3 + 2i$ con $5 + i$. Allora facciamo così:

$$(3 + 2i) + (5 + i) = (3 + 5) + (2 + 1)i = 8 + 3i.$$

Cioè sommiamo separatamente la parte "reale" e la parte con i. Altri esempi:

$$(-2 + 4i) + (7 - 3i) = 5 + i, \qquad (3 + 6i) - i = 3 + 7i, \qquad 3 + (-1 + i) = 2 + i.$$

Per i prodotti facciamo la stessa cosa: trattiamo i come una variabile formale e poi usiamo $i^2 = -1$:

$$(3 + 2i) \cdot (5 + i) = 3 \cdot 5 + 3 \cdot i + 2i \cdot 5 + 2i \cdot i = 15 + 3i + 10i + 2i^2 = 15 + 2i^2 + 13i.$$

Ora ricordiamo che $i^2 = -1$, quindi $2i^2 = -2$, e dunque

$$15 + 2i^2 + 13i = 15 - 2 + 13i = 13 + 13i.$$

Altri esempi:

$$(-2+4i)(7-3i) = -2+34i, \qquad (3+6i)\cdot(-i) = 6-3i, \qquad 3\cdot(-1+i) = -3+3i.$$

Ora definiamo il coniugato:

Definizione: Sia $a+bi$ un intero di Gauss. Definiamo il suo *coniugato* come $a-bi$.

Example 0.1.

- Il coniugato di $3+4i$ è $3-4i$.
- Il coniugato di $1-i$ è $1+i$.
- Il coniugato di 4 è 4.
- Il coniugato di $-6i$ è $6i$.

Osserviamo ora questo fatto utile. Se moltiplichiamo un intero di Gauss $a+bi$ per il suo coniugato $a-bi$, otteniamo

$$(a+bi)(a-bi) = a^2 - abi + abi - b^2i^2 = a^2 - b^2i^2$$

e siccome $i^2 = -1$ segue che

$$(a+bi)(a-bi) = a^2 + b^2.$$

Bello, no? Il nostro obiettivo è capire quando un primo è una somma di due quadrati, e ora abbiamo un modo semplice per produrre somme di due quadrati.

Che cos'è un primo negli interi di Gauss?

Se ricordate, nel secondo post abbiamo definito un primo p come un numero tale che, ogni volta che p divide un prodotto $a\cdot b$, allora p divide a oppure p divide b. Facciamo la stessa cosa negli interi di Gauss. Prima definiamo cosa significa "dividere" in $\mathbb{Z}[i]$.

Definizione: Sia $a+bi$ un intero di Gauss. Diciamo che $a+bi$ *divide* $c+di$ se possiamo scrivere

$$c+di = (a+bi)\cdot(e+fi)$$

per qualche intero di Gauss $e+fi$.

Definizione: Sia $a+bi$ un intero di Gauss. Diciamo che $a+bi$ è *primo* se, ogni volta che $a+bi$ divide un prodotto $(c+di)(e+fi)$, allora $a+bi$ divide $c+di$ oppure $a+bi$ divide $e+fi$.

Non dimostreremo il seguente teorema, la cui dimostrazione è simile a quella per gli interi.

Theorem 0.1 (Fattorizzazione unica)**.** *Ogni intero di Gauss $a+bi$ si può scrivere come prodotto di primi di Gauss, cioè*

$$a+bi = (c_1+d_1i)\cdots(c_k+d_ki)$$

per qualche intero positivo k. Questa scrittura è unica a meno di moltiplicare per $1, -1, i$ e $-i$.

Chiariamo questa unicità, che è un po' strana. Il punto è che, per esempio, l'intero di Gauss $1+3i$ si può fattorizzare come

$$1+3i = (1+i)(2+i).$$

(Potete verificare che $1+i$ e $2+i$ sono primi usando la definizione, ma può essere un po' laborioso.) Ora osserviamo che $i\cdot(-i) = -i^2 = 1$. Quindi possiamo scrivere

$$1+3i = 1\cdot(1+i)(2+i),$$

semplicemente moltiplicando per 1. Ma siccome $1 = i\cdot(-i)$, allora

$$1+3i = (i\cdot(-i))(1+i)(2+i),$$

che è lo stesso (per associatività) di

$$1+3i = i\cdot((-i)(1+i))(2+i),$$

e quindi

$$1+3i = i\cdot(1-i)(2+i).$$

Questa sembra un'altra fattorizzazione di $1+3i$. Tuttavia differisce dalla precedente solo perché abbiamo moltiplicato per $1 = i\cdot(-i)$. Theorem 0.1 ci sta dicendo che tutte le possibili fattorizzazioni in primi di $1+3i$ differiscono solo per moltiplicazioni "furbe" per 1. Per questo diciamo che, a meno di moltiplicare per questi elementi speciali (e se ci pensate si può fare anche scrivendo $1 = (-1)\cdot(-1)$ oppure $1 = (-i)\cdot i$), la fattorizzazione è unica.

Dimostrazione del teorema (finalmente!)

Ora possiamo capire come funzionerà la dimostrazione.

Sia p un numero primo negli interi. Uno schema della dimostrazione consiste nel provare le seguenti affermazioni:

- se $x^2+1 \equiv_p 0$ ha una soluzione allora $p+0i$ non è primo negli interi di Gauss;
- se $p+0i$ non è primo negli interi di Gauss, allora p si può scrivere come somma di due quadrati;
- se p è somma di due quadrati allora $p \equiv_4 1$ (lo abbiamo già dimostrato l'altra volta!);
- se $p \equiv_4 1$ allora $x^2+1 \equiv_p 0$ ha una soluzione.

Queste implicazioni ci dicono che, se sappiamo che una delle affermazioni sopra è vera, allora lo sono anche tutte le altre. In altre parole, le affermazioni

- $x^2 + 1 \equiv_p 0$ ha una soluzione;
- $p + 0i$ non è primo negli interi di Gauss;
- p è somma di due quadrati;
- $p \equiv_4 1$

stanno dicendo la stessa cosa: sono equivalenti, perché si implicano a vicenda.

Relazione tra $x^2 + 1 \equiv_p 0$ e il Teorema di Natale

Il seguente lemma è quello cruciale: grazie a questo lemma riusciremo a dimostrare il Teorema di Natale. Ci dà l'interpretazione corretta di cosa significa che un primo p è somma di due quadrati.

Lemma 0.1. *Sia p un numero primo. Se la congruenza*

$$x^2 + 1 \equiv_p 0$$

ha una soluzione, allora $p + 0i$ non è primo negli interi di Gauss.

Dire che $x^2 + 1 \equiv_p 0$ ha una soluzione significa che esiste un intero $x \in \mathbb{Z}$ tale che $x^2 + 1 \equiv_p 0$, oppure, in modo equivalente, che p divide $x^2 + 1$. Un altro modo di dirlo è

$$x^2 \equiv_p -1,$$

cioè stiamo dicendo che -1 è un quadrato modulo p.

Proof. Supponiamo che esista un $x \in \mathbb{Z}$ tale che

$$x^2 + 1 \equiv_p 0.$$

Questo significa che p divide $x^2 + 1$. Ma osserviamo che

$$x^2 + 1 = (x + i)(x - i)$$

negli interi di Gauss. Vogliamo mostrare che $p + 0i$ non è primo negli interi di Gauss.

Supponiamo per assurdo che $p + 0i$ sia primo negli interi di Gauss. Dato che $p + 0i$ divide il prodotto $(x + i)(x - i)$, allora dovrebbe dividere uno dei due fattori. Ma questo è impossibile. Infatti, se $p + 0i$ dividesse $x + i$, potremmo scrivere $x + i = (p + 0i)(a + bi) = pa + pbi$. Confrontando i coefficienti di i nei due membri otteniamo $1 = pb$, impossibile. Lo stesso argomento vale per $x - i$.

Quindi $p + 0i$ non è primo negli interi di Gauss. □

Primi negli interi che non sono primi negli interi di Gauss

Lemma 0.2. *Sia p un numero primo. Se $p+0i$ non è primo in $\mathbb{Z}[i]$, allora possiamo scrivere p come somma di due quadrati.*

Proof. Se $p+0i$ non è primo, allora (per Theorem 0.1) si può scrivere come prodotto di almeno due fattori:

$$p+0i = (a+bi)(c+di)$$

dove né $a+bi$ né $c+di$ sono uguali a $1, -1, i, -i$.

Ora moltiplichiamo entrambi i membri per i loro coniugati. A sinistra otteniamo

$$(p+0i)(p-0i) = p^2.$$

A destra otteniamo

$$(a+bi)(a-bi)(c+di)(c-di) = (a^2+b^2)(c^2+d^2).$$

Quindi

$$p^2 = (a^2+b^2)(c^2+d^2).$$

Abbiamo $a^2+b^2 > 1$ e $c^2+d^2 > 1$. Ma p è primo in $\mathbb{Z}$, quindi l'unico modo in cui due interi maggiori di 1 possono moltiplicarsi per dare p^2 è che

$$a^2+b^2 = p, \qquad c^2+d^2 = p.$$

In particolare,

$$p = a^2+b^2,$$

cioè p è una somma di due quadrati. □

Teorema di Wilson e soluzioni di $x^2+1 \equiv_p 0$

Useremo il seguente teorema per ottenere ciò che ci serve.

Theorem 0.1 (Teorema di Wilson)**.** *Se p è un numero primo, allora*

$$(p-1)\cdot(p-2)\cdot(p-3)\cdots 2\cdot 1 \equiv_p p-1.$$

Cioè, moltiplicando tutti i numeri da $p-1$ fino a 1 otteniamo $p-1$ modulo p.

Diamo un'idea della dimostrazione del teorema di Wilson; tuttavia questa dimostrazione non è essenziale per capire il cuore del Teorema di Natale.

Idea. Ricordiamo che un intero x è invertibile modulo p se e solo se $\text{MCD}(x, p) = 1$. Poiché p è primo, non ha fattori in comune con nessuno dei numeri $(p-1), (p-2), \ldots, 2, 1$. Questo significa che ogni fattore nel prodotto a sinistra è invertibile modulo p.

Ora, se elenchiamo tutti i resti invertibili modulo p, otteniamo esattamente $1, 2, \ldots, p-1$. Quindi nel prodotto a sinistra ogni termine compare insieme al suo inverso, tranne per i termini che sono il proprio inverso.

Un elemento a è il proprio inverso se e solo se $a^2 \equiv_p 1$. Modulo un primo questo significa $a \equiv_p 1$ oppure $a \equiv_p -1$. Quindi tutto si semplifica a coppie e rimane solo $-1 \equiv_p p-1$. □

Lemma 0.3. *Se p è un primo tale che $p \equiv_4 1$, allora $x^2 + 1 \equiv_p 0$ ha una soluzione.*

Proof. Poiché $p \equiv_4 1$, allora $p-1$ è divisibile per 4, cioè $p - 1 = 4k$ per qualche intero k. Consideriamo il prodotto

$$(p-1) \cdot (p-2) \cdot (p-3) \cdots 3 \cdot 2 \cdot 1.$$

Dato che $p - 1 = 4k$, possiamo raggruppare i fattori a coppie nel seguente modo:

$$(p-1) \cdot 1 \cdot (p-2) \cdot 2 \cdots (p-2k) \cdot (2k).$$

Poiché $p - j \equiv_p -j$, ogni coppia dà

$$(p-j) \cdot j \equiv_p (-j) \cdot j = -j^2.$$

Quindi

$$(p-1) \cdot (p-2) \cdots 2 \cdot 1 \equiv_p (-1)^2 \cdot (-2)^2 \cdots (-(2k))^2 = (1 \cdot 2 \cdot 3 \cdots \cdot (2k-1) \cdot 2k)^2.$$

(Qui i segni meno scompaiono perché stiamo elevando al quadrato.)

Ma per Theorem 0.1 sappiamo anche che il membro sinistro è congruo a $p-1$ modulo p. E poiché $p - 1 \equiv_p -1$, otteniamo

$$-1 \equiv_p (1 \cdot 2 \cdots (2k-1) \cdot 2k)^2.$$

Definiamo

$$x := 1 \cdot 2 \cdots (2k-1) \cdot 2k.$$

Allora

$$x^2 + 1 \equiv_p 0.$$

Quindi abbiamo trovato una soluzione. □

Il teorema è dimostrato: cos'altro possiamo dire?

Il teorema è finalmente dimostrato, ma non smettiamo di farci domande.

Domanda: Quanti primi si possono scrivere come somma di due quadrati? Equivalentemente, quanti primi hanno resto 1 nella divisione per 4?

Non sorprende che ci aspettiamo che esistano infiniti primi che si possono scrivere come somma di due quadrati. Questa affermazione è lasciata come esercizio guidato alla fine del post.

Un fatto interessante, ma non banale, è capire anche la densità di questo tipo di primi. Cioè: se consideriamo tutti i primi fino a un intero N, quanti ci aspettiamo che siano scrivibili come somma di due quadrati? La risposta è che ci aspettiamo circa la metà. Per esempio, questa tabella potrebbe convincervi:

N	primi $\leq N$	primi che sono a^2+b^2	proporzione
10^2	25	12	0.48
10^3	168	81	0.4821 ...
10^6	78498	39176	0.49907 ...
10^9	50847534	25423492	0.499994 ...

Questo si può dimostrare in modo rigoroso, ma è ben oltre i nostri scopi qui.

Un'altra domanda naturale potrebbe essere:

Domanda: Dato un intero n, quando lo si può scrivere come somma di due quadrati?

In altre parole, stiamo chiedendo come generalizzare dal caso dei primi al caso di un intero qualunque. Questo non è difficile da dimostrare una volta noto il Teorema di Natale, quindi lo lasciamo come esercizio guidato alla fine.

Conclusione

Vorrei sottolineare che il punto cruciale del nostro ragionamento è stato collegare il problema di scrivere p come somma di due quadrati con la risolubilità della congruenza x^2+1. Infatti, se ci pensate, gli interi di Gauss sono stati definiti aggiungendo agli interi una soluzione di $x^2+1=0$. Non è una coincidenza.

Esistono teoremi più generali che descrivono precisamente come la fattorizzazione di un primo (in realtà, di un ideale primo) in un'estensione degli interi (ottenuta aggiungendo qualcosa di simile alle soluzioni di un polinomio) sia legata alla fattorizzazione del polinomio modulo quel primo.

Nel nostro caso sapevamo che $x^2 + 1 \equiv_p 0$ ha una soluzione, cioè esiste un intero u tale che $u^2 \equiv_p -1$. Allora modulo p abbiamo la fattorizzazione

$$x^2 + 1 \equiv_p (x + u)(x - u),$$

perché

$$(x + u)(x - u) = x^2 - u^2 \equiv_p x^2 + 1.$$

Questo ci dice esattamente che il primo p si può scrivere come prodotto di due primi negli interi di Gauss (osservate la simmetria: il polinomio si spezza in due fattori esattamente come il primo). Il teorema che spiega questo comportamento è noto come teorema di Dedekind-Kummer.

Esercizi guidati

Esercizio 1: infiniti primi somma di due quadrati

Esercizio 1

Dimostra che esistono infiniti numeri primi che si possono scrivere come somma di due quadrati.

Suggerimento 1

Un primo è somma di due quadrati se e solo se -1 è un quadrato modulo p. Supponi che esistano solo un numero finito di primi tali che -1 sia un quadrato modulo p. Che cosa puoi dire?

Soluzione al Suggerimento 1

Dire "-1 è un quadrato modulo p" significa: esiste un intero n tale che

$$n^2 \equiv_p -1,$$

cioè $p \mid (n^2 + 1)$. Quindi ogni primo per cui -1 è un quadrato modulo p divide almeno un numero della forma $n^2 + 1$.

Suggerimento 2

Sia k il prodotto di tutti i primi p tali che -1 è un quadrato modulo p. Prova a imitare la dimostrazione di Euclide che esistono infiniti primi.

Suggerimento 3

Considera l'intero $k^2 + 1$. Quali primi possono dividerlo?

Soluzione al Suggerimento 3 (e fine dell'esercizio)

In modo analogo alla dimostrazione di Euclide, nessuno dei fattori primi di k (cioè i primi tali che -1 è un quadrato modulo p) può dividere k^2+1 (perché il resto sarebbe 1). Tuttavia deve esistere un primo q che divide k^2+1, quindi la lista finita che avevamo supposto non è completa.

Dunque l'ipotesi "solo un numero finito di primi con -1 quadrato modulo p" era falsa, ed esistono infiniti tali primi. Per il Teorema di Natale, questi sono esattamente i primi che sono somma di due quadrati.

Esercizio 2: quali interi sono somma di due quadrati?

Esercizio 2

Quali interi $n \geq 1$ si possono scrivere come $n = a^2 + b^2$?

Suggerimento 1

Che cosa succede se n è il prodotto di due primi che si possono scrivere come somma di due quadrati?

Soluzione al Suggerimento 1

Se $p = a^2 + b^2$ e $q = c^2 + d^2$, allora anche pq è una somma di due quadrati. Infatti una somma di due quadrati si può scrivere come un intero di Gauss per il suo coniugato:

$$a^2 + b^2 = (a + bi)(a - bi), \qquad c^2 + d^2 = (c + di)(c - di).$$

Quindi

$$(a^2 + b^2)(c^2 + d^2) = (a + bi)(a - bi)(c + di)(c - di).$$

Il fattore $(a + bi)(c + di)$ è un intero di Gauss, e $(a - bi)(c - di)$ è il suo coniugato. Quindi il loro prodotto è una somma di due quadrati.

Suggerimento 2

Considera il numero $5 \cdot 7^2$. Si può scrivere come somma di due quadrati?

Soluzione al Suggerimento 2

Sì:

$$5 \cdot 7^2 = 245 = 14^2 + 7^2.$$

(Quindi il primo $7 \equiv_4 3$ non è "vietato", ma sembra che conti l'esponente.)

Suggerimento 3

Nel suggerimento precedente, $5 \equiv_4 1$ e $7 \equiv_4 3$. Quale teorema ti aspetti?

Soluzione al Suggerimento 3 (enunciato)

Teorema (somma di due quadrati). Un intero positivo n si può scrivere come $n = a^2 + b^2$ se e solo se, nella fattorizzazione in primi di n, ogni primo $q \equiv_4 3$ compare con esponente **pari**.

Dimostrazione completa (guidata dai suggerimenti)

Passo 1: i primi $q \equiv_4 3$ devono comparire un numero pari di volte.

Dimostriamo prima il seguente lemma:

Lemma 0.1. *Sia q un primo con $q \equiv_4 3$. Se $q \mid (a^2 + b^2)$, allora $q \mid a$ e $q \mid b$.*

Proof. Supponiamo $q \mid (a^2 + b^2)$. Se $q \nmid b$, allora b ammette un inverso u modulo q e

$$(a \cdot u)^2 \equiv_q -1,$$

quindi -1 è un quadrato modulo q. Ma per il Teorema di Natale questo implicherebbe $q \equiv_4 1$, assurdo. Dunque $q \mid b$. Allora $q \mid a^2$ e quindi $q \mid a$. □

Ora supponiamo $n = a^2 + b^2$. Sia $q \equiv_4 3$ un primo che divide n. Per il lemma, q divide sia a sia b, quindi q^2 divide $a^2 + b^2 = n$. Questo significa che, nella fattorizzazione in primi di n, l'esponente di q non può essere dispari: deve essere pari.

Passo 2: il viceversa.

Ora supponiamo che n abbia la proprietà che ogni primo $q \equiv_4 3$ compaia con esponente pari. Allora possiamo scrivere n come prodotto di tre tipi di fattori:

- una potenza di 2;
- primi $p \equiv_4 1$ (eventualmente ripetuti);
- e quadrati di primi $q \equiv_4 3$ (eventualmente ripetuti).

  Ognuno di questi fattori è una somma di due quadrati:

- $2 = 1^2 + 1^2$;
- se $p \equiv_4 1$ allora p è somma di due quadrati per il Teorema di Natale;
- se hai un quadrato q^{2m} allora $q^{2m} = (q^m)^2 + 0^2$.

  Infine, per il Suggerimento 1, un prodotto di somme di due quadrati (e lo stesso ragionamento vale non solo per i primi) è ancora una somma di due quadrati. Quindi n è una somma di due quadrati.