VARIAZIONI GOLDBACH: PROBLEMI CON NUMERI PRIMI

Alessandro Zaccagnini

Conferenza tenuta per la Mostra "Oltre il Compasso." Parma, 21 ottobre 1998

Introduzione. Questo è il testo di una conferenza divulgativa tenuta in occasione dell'allestimento della mostra "Oltre il Compasso" a Parma nell'ottobre 1998. Se da un lato non si è ritenuto di modificare lo stile informale e colloquiale proprio di una conferenza a carattere divulgativo, è certamente opportuno premettere qualche parola alla stampa del testo. L'obiettivo della conferenza era quello di mostrare come anche argomentazioni del tutto elementari possano portare a formulare congetture profonde (tanto da essere ancora tali). In fondo, ogni teorema prima di essere dimostrato ha lo *status* di congettura nella mente del matematico che ne tenta la dimostrazione: e la storia insegna che può rimanere tale per secoli! Questo testo, dunque, non è didattico, strettamente parlando, né lo potrebbe essere data la sua origine: il suo eventuale interesse risiede nel tentativo di mostrare come parte della ricerca matematica, ivi inclusa quella ancora attuale, abbia almeno in parte la sua origine in problemi molto semplici da enunciare e sui quali è relativamente semplice fare congetture partendo da un livello di conoscenze matematiche elementare, ma aiutati da una certa dose di intuizione e dall'uso del ragionamento per analogia con casi già noti. Lo scopo è quello di invitare allo studio della matematica.

Preludio

Chiameremo *numeri primi* gli interi $n \geq 2$ che sono divisibili solo per 1 e per se stessi. Euclide (IV sec. a. C.) dimostrò per primo che esistono infiniti numeri primi. La sua dimostrazione è un eccellente esempio di ragionamento matematico: se 2, 3, 5, ..., p fossero i soli numeri primi, potremmo costruire il numero $2 \cdot 3 \cdot 5 \cdots p + 1$ il quale ha la proprietà di dare resto 1 quando diviso per ciascuno di questi numeri primi. Poiché ogni numero maggiore di 1 è primo oppure divisibile per almeno un numero primo, la nostra lista iniziale non può essere completa. Nelle parole del grande matematico inglese G. H. Hardy, questa dimostrazione "conserva la freschezza e l'importanza di quando è stata scoperta: duemila anni non vi hanno lasciato una ruga." Quello che segue è un piccolo elenco di numeri primi:

2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53	59	61	67
71	73	79	83	89	97	101	103	107	109	113	127	131	137	139	149	151	157	163
167	173	179	181	191	193	197	199	211	223	227	229	233	239	241	251	257	263	269
271	277	281	283	293	307	311	313	317	331	337	347	349	353	359	367	373	379	383
389	397	401	409	419	421	431	433	439	443	449	457	461	463	467	479	487	491	499
503	509	521	523	541	547	557	563	569	571	577	587	593	599	601	607	613	617	619
631	641	643	647	653	659	661	673	677	683	691	701	709	719	727	733	739	743	751
757	761	769	773	787	797	809	811	821	823	827	829	839	853	857	859	863	877	881
883	887	907	911	919	929	937	941	947	953	967	971	977	983	991	997			

Tavola 1. I numeri primi fino a 1000.

© L'Educazione Matematica, Anno XXI, Serie VI **2** (2000), 47–57. Riprodotto per gentile concessione.

Ora che sappiamo che esistono infiniti numeri primi, possiamo chiederci se esista un metodo per determinarli, a parte la divisione per tentativi. Eratostene (II sec. a. C.) inventò il cosiddetto crivello (cioè setaccio) che permette di separare i numeri primi dai numeri composti: la Tavola 2 mostra come si procede.

1	2	3	4	5	6	7	8	9	10	11	12
13	14	15	16	17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32	33	34	35	36
37	38	39	40	41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70	71	72
73	74	75	76	77	78	79	80	81	82	83	84
85	86	87	88	89	90	91	92	93	94	95	96
97	98	99	100	101	102	103	104	105	106	107	108
109	110	111	112	113	114	115	116	117	118	119	120
121	122	123	124	125	126	127	128	129	130	131	132
133	134	135	136	137	138	139	140	141	142	143	144

Tavola 2. Il crivello di Eratostene.

Il crivello funziona cosí: a parte il numero 1, che ha uno *status* speciale, cancelliamo dallo schema qui sopra tutti i multipli di 2 a partire da $2^2 = 4$. Poi guardiamo qual è il piú piccolo numero non cancellato, 3, e procediamo come prima, partendo da $3^2 = 9$. Ripetiamo queste operazioni con 5, a partire da $5^2 = 25$, poi con 7, partendo da $7^2 = 49$, ed infine con 11, partendo da $11^2 = 121$. A questo punto possiamo fermarci, poiché il primo numero non ancora cancellato è 13, e $13^2 = 169$ che è fuori dalla nostra tavola. La nostra tavola mostra 1 e tutti i numeri primi fino a 144. Le righe aiutano a cancellare i multipli dello stesso numero primo.

Per quanto dimostrato da Euclide i numeri primi non finiscono mai, ma osservando attentamente la Tavola 1 si vede che la loro densità sembra diminuire lentamente, e questo può essere confermato costruendo tavole piú grandi. Possiamo spiegare questo fenomeno? Eseguendo le operazioni richieste dal crivello sugli interi $1, 2, \dots, N$, (dove N è un numero molto grande) possiamo notare che solo la metà circa degli interi sopravvivono al primo passaggio (quello relativo al numero primo 2), e dei rimanenti solo i $\frac{2}{3}$ rimangono dopo il secondo passaggio ($p = 3$), e cosí via. In altre parole, quando usiamo il numero primo p cancelliamo circa $1/p$-esimo degli interi nella nostra tavola che non sono stati ancora cancellati. Dato che ogni intero non primo ha almeno un fattore primo che non supera la sua radice quadrata, alla fine la proporzione dei numeri rimasti nella tavola sul totale dovrebbe essere approssimativamente

$$\left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{3}\right) \cdot \left(1 - \frac{1}{5}\right) \cdots \left(1 - \frac{1}{p}\right),$$

dove abbiamo indicato con p il piú grande numero primo $\leq \sqrt{N}$.

Si può dimostrare, anche se in questa sede non è possibile neppure accennare come, che il prodotto scritto sopra vale approssimativamente

$$\frac{c}{\log N} \tag{1}$$

dove $c = 1.1229197\dots$, se N è un numero molto grande. Qui e nel seguito log indica la funzione *logaritmo naturale*, o in base $e = 2.718281828\dots$ A questo punto sembrerebbe legittimo aspettarsi, sempre se N è un numero molto grande, che debbano esserci approssimativamente $cN/\log N$ numeri primi $\le N$, dato che (1) sembra rappresentare la proporzione di numeri non cancellati nel crivello di Eratostene. Le cose non stanno esattamente cosí, perché per N grande la proporzione dei numeri cancellati dal primo $p \le \sqrt{N}$ non è quasi mai esattamente $1/p$: questa cosa accade solo in casi eccezionali. I passaggi che portano alla (1) non sono dunque pienamente giustificabili, ma con un'argomentazione diversa (piuttosto lunga e intricata, e quindi non ne parleremo), nel 1896 Jacques Hadamard e Charles de la Vallée Poussin hanno dimostrato indipendentemente che

$$\pi(N) \sim \frac{N}{\log N}, \tag{2}$$

dove tradizionalmente $\pi(N)$ indica il numero dei numeri primi che non superano N. Non c'è pericolo di confusione con l'altro significato di π: in questa conferenza il numero $\pi = 3.14159\ldots$ non comparirà mai. Il simbolo $\sim$ indica un'uguaglianza approssimata: piú precisamente, la (2) significa che il rapporto fra le due quantità $\pi(N)$ ed $N/\log N$ è molto vicino ad 1 quando N è molto grande. Questa relazione è nota come Teorema dei Numeri Primi, ed è il frutto di oltre un secolo di sforzi di matematici di prima grandezza come L. Eulero, C. F. Gauss, P. L. Dirichlet, B. Riemann, P. Chebyshev.

La Tavola 3 mostra il valore di $\pi(N)$ quando N è una potenza di 10 piccola, l'errore commesso nell'approssimazione ed il rapporto fra le quantità che compaiono nella (2).

N	$\pi(N)$	$\pi(N) - \dfrac{N}{\log N}$	$\dfrac{\pi(N)\log N}{N}$
10	4	0	$0.921\ldots$
10^2	25	3	$1.151\ldots$
10^3	168	23	$1.161\ldots$
10^4	1229	143	$1.132\ldots$
10^5	9592	906	$1.104\ldots$
10^6	78498	6116	$1.084\ldots$
10^7	664579	44158	$1.071\ldots$
10^8	5761455	332774	$1.061\ldots$
10^9	50847534	2592592	$1.054\ldots$
10^{10}	455052511	20758029	$1.048\ldots$

Tavola 3. Valori di $\pi(N)$ per $N = 10^n$, $n = 1, \ldots, 10$. Le differenze nella terza colonna sono approssimate all'intero piú vicino.

Tema: il problema di Goldbach

Nel 1742 il matematico Christian Goldbach affermò in una lettera ad Eulero che ogni numero intero pari maggiore di 4 può essere scritto come somma di due numeri primi dispari (non necessariamente distinti). In altre parole, se n è un numero intero pari maggiore di 4 è possibile trovare due numeri primi dispari p_1 e p_2 in modo che

$$n = p_1 + p_2. \tag{3}$$

Per esempio, $10 = 3+7 = 5+5 = 7+3$. Questo oggi è noto con il nome di "problema binario di Goldbach," e non se ne è ancora trovata una dimostrazione. Nella Tavola 4 troverete il numero delle soluzioni dell'equazione (3) (indicato con $r(n)$) per n pari fino a 200. $r(4) = 1$

perché $4 = 2 + 2$, e inoltre, come abbiamo fatto sopra con 10, le soluzioni come $3 + 7$ e $7 + 3$ sono considerate distinte.

n	$r(n)$	n	$r(n)$	n	$r(n)$	n	$r(n)$	n	$r(n)$
2	0	4	1	6	1	8	2	10	3
12	2	14	3	16	4	18	4	20	4
22	5	24	6	26	5	28	4	30	6
32	4	34	7	36	8	38	3	40	6
42	8	44	6	46	7	48	10	50	8
52	6	54	10	56	6	58	7	60	12
62	5	64	10	66	12	68	4	70	10
72	12	74	9	76	10	78	14	80	8
82	9	84	16	86	9	88	8	90	18
92	8	94	9	96	14	98	6	100	12
102	16	104	10	106	11	108	16	110	12
112	14	114	20	116	12	118	11	120	24
122	7	124	10	126	20	128	6	130	14
132	18	134	11	136	10	138	16	140	14
142	15	144	22	146	11	148	10	150	24
152	8	154	16	156	22	158	9	160	16
162	20	164	10	166	11	168	26	170	18
172	12	174	22	176	14	178	13	180	28
182	12	184	16	186	26	188	10	190	16
192	22	194	13	196	18	198	26	200	16

Tavola 4. Valori di $r(n)$ per n pari fino a 200.

In questo caso, un'analisi della Tavola 4 sembra mostrare che il numero delle soluzioni $r(n)$ cresce quando n cresce, ma in modo estremamente irregolare: per esempio, $r(180) = 28$, mentre $r(182) = 12$. Possiamo provare a spiegare anche questo fenomeno?

Prima di tutto, è abbastanza ragionevole che $r(n)$ debba crescere in qualche modo, dato che piú grande è n, piú sono numerosi i numeri primi che possono comparire come addendi nella relazione (3). Inoltre, possiamo servirci della (2) per ottenere un "ordine di grandezza" atteso per $r(n)$. Preso un numero N molto grande, consideriamo i $\pi(N) - 1$ numeri primi dispari minori o uguali ad N, e tutte le loro possibili somme. Naturalmente, se $p_1 \le N$ e $p_2 \le N$ possiamo solo concludere che $p_1 + p_2 \le 2N$, ma poco importa, dato che ci accontentiamo di un risultato approssimato. Per la (2) le possibili somme $p_1 + p_2$ sono

$$\bigl(\pi(N) - 1\bigr)^2 \sim \frac{N^2}{(\log N)^2},$$

e quindi, *in media*, ogni intero pari $n \le 2N$ ha approssimativamente $N/(\log N)^2$ rappresentazioni del tipo (3) con p_1 e $p_2 \le N$, poiché ci sono esattamente N interi pari $\le 2N$. Questo ragionamento in media non spiega ancora le pronunciate irregolarità della Tavola 4, ma suggerisce che $r(n)$ debba essere vicino a

$$\frac{n}{(\log n)^2}. \tag{4}$$

Come possiamo dare una spiegazione convincente di quello che osserviamo? In altre parole, la (4) dà veramente il giusto ordine di grandezza per $r(n)$? E se la risposta è affermativa, come si spiegano le irregolarità di $r(n)$? Nella Tavola 5 abbiamo riportato le soluzioni dell'equazione $a+b=n$ per $n=60$ e per $n=62$, dove a e b sono in ogni caso numeri interi dispari ed $a \leq b$ (queste semplificazioni inessenziali servono per evitare una tavola enorme).

$1 + 59$	$1 + 61$
$\ast\ 3 + 57\ \ast$	$\ast\ 3 + 59$
$5 + 55$	$5 + 57\ \ast$
$7 + 53$	$7 + 55$
$\ast\ 9 + 51\ \ast$	$\ast\ 9 + 53$
$11 + 49$	$11 + 51\ \ast$
$13 + 47$	$13 + 49$
$\ast\ 15 + 45\ \ast$	$\ast\ 15 + 47$
$17 + 43$	$17 + 45\ \ast$
$19 + 41$	$19 + 43$
$\ast\ 21 + 39\ \ast$	$\ast\ 21 + 41$
$23 + 37$	$23 + 39\ \ast$
$25 + 35$	$25 + 37$
$\ast\ 27 + 33\ \ast$	$\ast\ 27 + 35$
$29 + 31$	$29 + 33\ \ast$
	$31 + 31$

Tavola 5. Come ottenere le soluzioni dell'equazione (3) quando $n=60$ oppure $n=62$. Gli asterischi indicano i numeri divisibili per 3.

Ora operiamo una specie di doppio crivello, cancellando dalle potenziali soluzioni dell'equazione (3) quelle in cui un addendo risulta divisibile per 3, indicate dagli asterischi. Si nota immediatamente una differenza importante fra le due situazioni: nel caso di $n=60$ gli asterischi compaiono alla stessa altezza, e quindi quando consideriamo il numero primo 3 cancelliamo solo $1/3$ delle soluzioni dell'equazione $a+b=60$ elencate all'inizio, e ne sopravvivono $1-1/3=2/3$. Invece, quando $n=62$ cancelliamo $2/3$ delle soluzioni dell'equazione $a+b=62$, lasciandone solamente $1-2/3=1/3$.

Come è possibile distinguere fra le due situazioni? Non è molto difficile vedere che la differenza risiede nel fatto che 3 divide 60 ma non 62. Naturalmente il numero primo 3 non ha niente di speciale: lo stesso ragionamento vale per tutti i numeri primi dispari $p \leq \sqrt{n}$. Quindi, procedendo in modo analogo a prima, siamo tentati di modificare la precedente congettura che $r(n)$ sia circa la quantità (4), sostituendola con

$$n \cdot \frac{1}{2} \cdot \left(1-\frac{1}{p_1}\right) \cdot \left(1-\frac{1}{p_2}\right) \cdots \left(1-\frac{1}{p_r}\right) \cdot \left(1-\frac{2}{q_1}\right) \cdot \left(1-\frac{2}{q_2}\right) \cdots \left(1-\frac{2}{q_s}\right), \tag{5}$$

dove $p_1, \ldots, p_r$ sono i primi dispari distinti che dividono n, mentre $q_1, \ldots, q_s$ sono gli altri primi dispari minori di $\sqrt{n}$, ed il fattore $\frac{1}{2}$ tiene conto del fatto che non ci possono essere addendi pari nella (3) se $n > 4$.

Per fortuna non è difficile semplificare un po' l'espressione (5): tenendo anche conto delle differenze fra la (1) e la (2), dopo qualche calcolo la nostra congettura prende la forma

$$r(n) \sim c' \cdot \frac{p_1-1}{p_1-2} \cdot \frac{p_2-1}{p_2-2} \cdots \frac{p_r-1}{p_r-2} \cdot \frac{n}{(\log n)^2}, \tag{6}$$

per un'opportuna costante $c' = 1.3203\ldots$ detta "costante dei primi gemelli" per un motivo che vedremo fra breve. Osserviamo che, per esempio, quando $n=60$ il prodotto sui fattori primi nella (6) vale $8/3 = 2.6666\ldots$ mentre per $n = 62$ vale solamente $30/29 = 1.03448\ldots$ Dunque, se il ragionamento che abbiamo fatto è corretto, non è piú un mistero il motivo per cui $r(60) = 12$ è piú del doppio di $r(62) = 5$.

Abbiamo detto sopra che il problema di Goldbach non è stato ancora risolto: ci sono altre argomentazioni di natura completamente diversa che suggeriscono la correttezza della formula (6), che rappresenta una forma molto forte della congettura di Goldbach, ma finora nessuno è stato in grado di fornirne una dimostrazione rigorosa. In effetti questa congettura è stata verificata per mezzo di calcolatori fino ad $n = 4 \cdot 10^{14}$, ma questa non può certo essere considerata una dimostrazione. Una parte della difficoltà risiede nel fatto che l'approssimazione data dalla formula (2) non è sufficientemente buona. In ogni caso, sono stati dimostrati dei risultati "approssimati" che in questa sede è difficile descrivere. Il piú semplice da raccontare fra questi è oggetto della ...

Prima variazione

Nella (3) abbiamo considerato il problema di rappresentare i numeri pari abbastanza grandi (qui abbastanza grandi significa maggiori di 4) nella forma $p_1 + p_2$, dove p_1 e p_2 sono primi dispari. Ora ci chiediamo: qual è il problema corrispondente per i numeri interi dispari?

Poiché tutti i numeri primi a parte 2 sono dispari, dobbiamo sommare tre numeri primi (dispari) per ottenere un altro numero dispari. Dunque, chiameremo "congettura ternaria di Goldbach" l'affermazione seguente: comunque dato un numero dispari n sufficientemente grande, esistono tre numeri primi dispari p_1, p_2 e p_3, non necessariamente distinti, tali che

$$n = p_1 + p_2 + p_3. \tag{7}$$

Naturalmente è possibile di nuovo ragionare per analogia con i casi precedenti, operando con un crivello triplo. Inoltre ci si può aspettare che questo problema sia piú facile da risolvere del precedente perché il ragionamento in media che portava alla (4) ora dà il numero molto piú grande $n^2/(\log n)^3$. In effetti, nel 1937 il matematico russo I. M. Vinogradov dimostrò che per n dispari sufficientemente grande l'equazione (7) ha sempre almeno una soluzione. In piú Vinogradov dimostrò anche una formula simile alla (6) per il numero delle soluzioni, che chiameremo $r_3(n)$; se n è un numero dispari

$$r_3(n) \sim c'' \cdot \frac{(p_1-1)(p_1-2)}{p_1^2 - 3p_1 + 3} \cdot \frac{(p_2-1)(p_2-2)}{p_2^2 - 3p_2 + 3} \cdots \frac{(p_r-1)(p_r-2)}{p_r^2 - 3p_r + 3} \cdot \frac{n^2}{(\log n)^3}, \tag{8}$$

dove c'' è un'altra costante positiva, e p_1, p_2, ... , p_r sono i fattori primi di n.

All'epoca Vinogradov non riuscí a precisare cosa volesse dire "sufficientemente grande," ma studi successivi dimostrarono che ogni $n > 3^{3^{15}}$ è sufficientemente grande. Purtroppo questo è un numero enorme (ha quasi 7 milioni di cifre) che è stato poi ridotto, ma non è ancora possibile trattare i casi rimanenti con un calcolatore. Recentemente il matematico francese O. Ramaré ha dimostrato che *qualunque* sia il numero naturale $n \geq 2$, l'equazione

$$n = p_1 + p_2 + \cdots + p_r$$

è risolubile con i p numeri primi, ed $r \leq 7$.

Seconda variazione

Proviamo a modificare l'equazione (3) cambiando il segno $+$ in un segno $-$:

$$n = p_1 - p_2. \tag{9}$$

La conseguenza piú importante di questa modifica è che, mentre l'equazione (3) poteva avere solo un numero finito di soluzioni, per l'equazione (9) non è detto che debba essere cosí. Per poter confrontare le due situazioni, consideriamo un numero grande N e contiamo le soluzioni dell'equazione (9) in cui $p_2 \le N$. La Tavola 6 dà il numero delle soluzioni dell'equazione (9) (indicate con $\pi_n(N)$) per n pari fino a 100 ed $N = 10000$. Si nota subito che tutti i numeri nella Tavola sono piuttosto grandi, ma sono evidenti delle irregolarità: tornando al nostro esempio precedente, $\pi_{60}(10000)$ è oltre il doppio di $\pi_{62}(10000)$. Ci chiediamo: è possibile trovare una semplice spiegazione di questo fenomeno?

n	π_n	n	π_n	n	π_n	n	π_n	n	π_n
2	205	4	203	6	411	8	208	10	270
12	404	14	245	16	200	18	417	20	269
22	226	24	404	26	240	28	248	30	536
32	196	34	215	36	404	38	213	40	267
42	489	44	227	46	201	48	409	50	270
52	221	54	410	56	240	58	212	60	535
62	206	64	201	66	458	68	209	70	318
72	401	74	206	76	220	78	428	80	272
82	205	84	493	86	207	88	217	90	531
92	218	94	208	96	400	98	232	100	260

Tavola 6. Valori di $\pi_n(10000)$ per n pari fino a 100.

Una volta fissato $n = 60$ ed $N = 10000$, per esempio, possiamo costruire l'analoga della Tavola 5, in cui riportiamo nella prima colonna i numeri dispari 1, 3, ... , 9999, e nella seconda colonna i numeri $1 + 60$, $3 + 60$, ... , $9999 + 60$. Eliminando come prima i numeri divisibili per 3 da ciascuna colonna, vediamo che anche in questo caso essi appaiono alla stessa altezza, mentre nel procedimento relativo ad $n = 62$ sono ad altezze diverse. Siamo quindi condotti a congetturare che per $\pi_n(N)$ debba valere una relazione quale

$$\pi_n(N) \sim c' \cdot \frac{p_1 - 1}{p_1 - 2} \cdot \frac{p_2 - 1}{p_2 - 2} \cdots \frac{p_r - 1}{p_r - 2} \cdot \frac{N}{(\log N)^2}, \tag{10}$$

dove, come prima, p_1, ... , p_r sono i fattori primi dispari di n (che consideriamo fissato), N è un numero molto grande, e c' è la stessa costante che compariva in (6). Possiamo finalmente spiegare il nome di questa costante. In origine il problema è stato formulato cosí: è vero che esistono infiniti "primi gemelli," primi cioè che come 11 e 13 hanno differenza 2? Qui abbiamo preferito trattare insieme tutti i casi relativi alle differenze pari. Concludiamo osservando che i valori piú piccoli nella Tavola 6 sono quelli per cui n è una potenza di 2, mentre i piú grandi sono quelli per cui n ha molti fattori primi dispari piccoli, come 6, 30, ... , in buon accordo con la (10). Anche questo, come quello di Goldbach, è un problema aperto, per il quale sono noti risultati "approssimati" difficili da descrivere in questa sede.

Altre variazioni

Si diceva che il problema dei primi gemelli può essere formulato in un modo leggermente diverso, che suggerisce una nuova variazione: se davvero esistono infiniti numeri primi p per cui anche $p+2$ è un numero primo, sarà vero che esistono infiniti numeri primi p per cui anche $p + 2$ e $p + 4$ sono numeri primi? Guardando la Tavola 1 scopriamo che questo accade una sola volta, quando $p = 3$. Non è difficile capire dov'è il problema: qualunque sia il numero intero n, uno fra i numeri n, $n+2$ ed $n+4$ è divisibile per 3. Ma se modifichiamo la richiesta, e guardiamo ai numeri p, $p+2$ e $p+6$, scopriamo che questi sono simultaneamente primi in una grande quantità di casi: in effetti, esistono ben 55 numeri primi $p < 10000$ tali che $p+2$ e $p+6$ sono ancora numeri primi.

Siamo quindi condotti a considerare quelle che alcuni matematici chiamano *costellazioni* di numeri primi: dati k numeri interi positivi distinti a_1, a_2, ... , a_k, ci chiediamo se esistano infiniti primi p tali che anche $p+a_1$, ... , $p+a_k$ siano tutti simultaneamente primi. Evidentemente il caso dei primi gemelli corrisponde a prendere $k=1$ ed $a_1=2$. Per semplificare il ragionamento che segue poniamo $a_0=0$ e scriviamo in forma compatta $(a_0,\dots,a_k)$ per indicare i $k+1$ numeri a_0, a_1, ... , a_k e quindi per identificare la costellazione che ci interessa. Il problema che ci si pone immediatamente davanti è come distinguere fra la situazione in cui, per esempio, $k=2$, $a_1=2$ ed $a_2=4$ (cioè la costellazione $(0,2,4)$) e quella in cui $k=2$, $a_1=2$ ed $a_2=6$ (cioè la costellazione $(0,2,6)$), che come abbiamo visto sopra sono radicalmente diverse.

Si può procedere in questo modo: consideriamo i numeri primi p che non superano $k+1$ ($k+1$ è il numero dei numeri primi nella costellazione) e per ciascuno di questi numeri primi calcoliamo il resto di a_0, a_1, ... , a_k nella divisione per p. Se per qualcuno di questi primi accade che i resti, in un qualche ordine, sono 0, 1, ... , $p-1$, allora i numeri $p=p+a_0$, $p+a_1$, ... , $p+a_k$ possono essere simultaneamente primi al massimo per un valore di p, e la costellazione $(a_0,\dots,a_k)$ viene detta *non ammissibile*. Viceversa, se per ogni primo $\le k+1$ almeno un resto non compare, la costellazione si dice *ammissibile*.

Per tornare al nostro esempio di prima, nel caso $(0,2,4)$ vediamo immediatamente che se $p=2$ i resti sono sempre 0, mentre se $p=3$ i resti valgono 0, 2, 1 rispettivamente. Per $(0,2,6)$ e $p=2$ tutti i resti valgono 0, ma per $p=3$ i resti sono 0, 2, 0.

Le idee esposte sopra permettono di congetturare che, data una costellazione ammissibile $(a_0,\dots,a_k)$, il numero di primi p per cui anche $p+a_1$, ... , $p+a_k$ sono numeri primi sia infinito, ma la formula corrispondente alla (10) è piuttosto complicata. Ma perché limitarci a considerare soltanto costellazioni di primi, a distanza per cosí dire fissata? In effetti, questo stesso metodo può essere utilizzato per studiare il problema della distribuzione dei numeri primi p per cui anche $2p+1$ è un numero primo, oppure potremmo scegliere relazioni piú complicate. Questa considerazione suggerisce che ci siano infiniti problemi di questo tipo, ma preferiamo concludere qui perché il tempo a disposizione è finito.

a	b	$\pi_{(0,a,b)}$	a	b	$\pi_{(0,a,b)}$	a	b	$\pi_{(0,a,b)}$	a	b	$\pi_{(0,a,b)}$
2	6	55	2	8	57	2	12	92	2	14	73
2	18	66	2	20	82	2	24	59	2	26	68
2	30	112	2	32	85	2	36	65	4	6	57
4	10	91	4	12	62	4	16	57	4	18	71
6	8	62	6	12	118	6	14	76	6	16	81
6	18	125	6	20	106	6	22	62	8	12	49
8	14	68	8	18	88	12	60	166	20	30	101
30	60	223	30	90	221	30	120	219	60	120	228

Tavola 7. Conteggi per alcune costellazioni di primi $(0,a,b)$, con $N=10000$. Sono state prese in considerazione solo costellazioni ammissibili.

Coda

Queste ultime parole sono per quelle persone che insistono per sapere tutta la verità. La costante c nella (1) vale $2e^{-\gamma}$, dove γ è la costante di Eulero-Mascheroni, definita da

$$\gamma = \lim_{n\to\infty}\left(\sum_{i=1}^{n}\frac{1}{i} - \log n\right) = 0.5772156649\ldots$$

La dimostrazione non è semplice. È possibile migliorare l'approssimazione a $\pi(N)$ data dalla (2), sostituendo $N/\log N$ con $N/(\log N - 1)$, o meglio ancora con la funzione *logaritmo integrale*, definita da

$$\mathrm{li}(N) = \int_2^N \frac{\mathrm{d}t}{\log t}.$$

Tutto questo può essere spiegato dalla teoria generale. Con quest'ultima funzione, la differenza $\pi(N) - \mathrm{li}(N)$ risulta molto piú piccola di quella nella terza colonna della Tavola 3. Si congettura che esista una costante positiva A tale che per N grande si abbia

$$\left|\pi(N) - \mathrm{li}(N)\right| \le A\sqrt{N}\log N,$$

ma le migliori stime note al giorno d'oggi sono molto piú deboli.

È stato dimostrato che sia $r(n)$ che $\pi_n(N)$ non superano il quadruplo dell'espressione a destra nella (6) e nella (10) rispettivamente, e in generale tutte le quantità di cui abbiamo parlato non superano un multiplo fissato del numero atteso. Inoltre gli interi pari n per cui $r(n)$ è significativamente piú piccolo della quantità (6) sono rari. L'enunciato preciso è piuttosto tecnico.

Hardy & Littlewood nel loro famoso articolo [3] fanno la storia della congettura (6), e dimostrano che formule analoghe congetturate in precedenza (che, essenzialmente, avevano una diversa costante al posto di c' oppure un prodotto sui fattori primi dispari di n di forma diversa) sono certamente false. Il loro articolo contiene un'accurata critica dei metodi "probabilistici" che conducono a queste formule sbagliate, i quali utilizzano ragionamenti simili a quello che porta alla (1), ed anche molte tavole numeriche. Inoltre, nello stesso articolo hanno descritto un metodo rivoluzionario per affrontare questi e simili problemi, e quasi tutti i risultati piú forti ottenuti da allora (compreso quello di Vinogradov) sono stati dimostrati utilizzando il loro metodo. Non è possibile accennare neppure vagamente alle idee utilizzate: sono necessarie profonde conoscenze di analisi reale e complessa.

Le costanti c' e c'' valgono rispettivamente

$$c' = 2\prod_{p\neq 2}\left(1 - \frac{1}{(p-1)^2}\right) \qquad \text{e} \qquad c'' = \prod_{p\neq 2}\left(1 + \frac{1}{(p-1)^3}\right),$$

dove i prodotti sono estesi a tutti i numeri primi dispari. Il numero $3^{3^{15}}$ è stato portato ad e^{100000}, ma anche questo numero è molto grande (ha oltre 43000 cifre).

Per quello che riguarda le costellazioni di primi, chiamiamo $\rho(p)$ il numero dei resti *distinti* dei numeri $a_0, \ldots, a_k$ nella divisione per p. Non è difficile dimostrare che possiamo riformulare il ragionamento fatto sopra dicendo che la costellazione $(a_0, \ldots, a_k)$ è ammissibile se e soltanto se $\rho(p) < p$ per ogni numero primo p. Il numero atteso di numeri primi $p \le N$ per cui anche $p + a_1, \ldots, p + a_k$ sono numeri primi è dato dalla formula

$$\pi_{(a_0,a_1,\ldots,a_k)}(N) \sim \prod_p \left(1 - \frac{1}{p}\right)^{-1-k}\left(1 - \frac{\rho(p)}{p}\right)\cdot\frac{N}{(\log N)^{k+1}},$$

che contiene la (2) e la (10) come casi particolari. Per esempio, nel caso $k = 2$ illustrato dalla Tavola 7 si ha $\rho(p) = 3$ per tutti i numeri primi che non dividono $ab(b-a)$.

Nota. Questa conferenza deve molto alla prima di quelle contenute nel libro di Serge Lang [5], mentre il titolo è ispirato a quello del dialogo "Variazioni Goldbach" nel libro di D. R. Hofstadter [4]. Inoltre i dati della Tavola 3 sono elaborati a partire da quelli contenuti nelle Tavole 5.2 e 26 rispettivamente dei libri di Conway & Guy [1], e di Ribenboim [7], mentre gli altri dati sono stati calcolati dall'autore. Per ulteriori informazioni si vedano i riferimenti bibliografici e le pagine web qui sotto:

`http://www.inesca.pt/~tos/goldbach.html`
Risultati della verifica della Congettura di Goldbach di Tomás Oliveira e Silva.

`http://www.informatik.uni-giessen.de/staff/richstein/ca/Goldbach.html`
Verifica della Congettura di Goldbach fino a $4 \cdot 10^{14}$ di Jörg Richstein.

`http://purl.oclc.org/NET/TRN/`
Calcolo dei primi gemelli fino a 10^{14} di Thomas Nicely.

Ringrazio Keith Matthews per il riferimento a queste pagine. Questo testo è disponibile su rete in formato pdf all'indirizzo
`http://www.math.unipr.it/~zaccagni/psfiles/Goldbach_I.pdf`
ed in inglese ad
`http://www.math.unipr.it/~zaccagni/psfiles/Goldbach_E.pdf`

Bibliografia

[1] J. H. Conway & R. K. Guy, *Il Libro dei Numeri*, Hoepli, Milano, 1999.
[2] G. H. Hardy, *Apologia di un Matematico*, Garzanti, Milano, 1989.
[3] G. H. Hardy & J. E. Littlewood, *Some problems of "Partitio Numerorum"; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.
[4] D. R. Hofstadter, *Gödel, Escher, Bach: un'Eterna Ghirlanda Brillante*, Adelphi, Milano, 1984.
[5] S. Lang, *La Bellezza della Matematica*, Bollati Boringhieri, 1991.
[6] C. Pomerance, *Alla ricerca dei numeri primi*, Le Scienze **174** (febb. 1983), 86–94.
[7] P. Ribenboim, *The New Book of Prime Number Records*, Springer, Berlino, 1996.

Dipartimento di Matematica, Università di Parma, via Massimo d'Azeglio 85/A, 43100, Parma
E-mail address: `zaccagnini@prmat.math.unipr.it`
URL: `http://www.math.unipr.it/~zaccagni/home.html`