整数数列中素因子个数的问题

付艺渲

(山东省实验中学, 济南, 250001)

站在数论的角度研究一个整数数列, 其中是否有无穷多个素数显然是让人极感兴趣的性质. 但我们甚至不能知道简单的数列, 如 n^2+1 中是否有无穷多个素数. 退一步, 一个整数数列中是否有无穷多个素因子, 也可以算是它的基本的数论性质. 在数学竞赛中, 不乏这类问题. 本文主要介绍这一方面的结论与问题.

我们首先来看一下有关的基本定理.

定理 1 (Issai Schur[1]). 设 $f(x)$ 是一个非常数的整系数多项式, 则数列 $\{f(n)\}$ $(n=0,1,\cdots)$ 有无穷多个不同的素因数.

证明 设非常数的整系数多项式 $f(x)=\sum\limits_{i=0}^{n}a_ix^i$ 对 $x=0,1,\cdots$ 仅有有限个不同的素因子 $p_1,\ldots,p_k$, 则 $a_0\neq 0$.

取 $x=p_1\cdots p_ka_0t$, 并设整数 t 充分大, 则 $f(x)$ 可表示为如下形式:

$$f(x)=a_0(p_1\cdots p_kA_t+1),$$

这里 A_t 是一个依赖于 t 的整数. 在 t 充分大时, $|A_t|>1$, 故 $p_1\cdots p_kA_t+1$ 有素因子 p, 显然 p 不同于 $p_1,\ldots,p_k$, 与假设矛盾. □

评注 定理 1 的应用无疑是极其广泛的. 这里的证明手法, 与 Euclid 证明素数无限的方式如出一辙.

定理 2. 若 f 是一个首项系数大于 0 的整系数多项式, $\{a_n\}$ 是一个严格递增的正整数数列, 且对于任意正整数 n, 都有 $a_n\leq f(n)$, 则 $\{a_n\}$ 中有无穷多个不同的素因数.

证明 若结论不成立, 则只存在有限个素数 $p_1,\ldots,p_t$ 满足要求.

设多项式 f 的首项为 b_mx^m, 由于每个 a_n 都可以写成 $p_1,\ldots,p_t$ 的幂次之积,

收稿日期: 2018-02-14;　修订日期: 2018-04-29.

因此
$$\sum_{n=1}^{\infty}\frac{1}{f(n)^{\frac{1}{m}}}\le\sum_{n=1}^{\infty}\frac{1}{a_n^{\frac{1}{m}}}\le\prod_{i=1}^{t}\sum_{j=1}^{\infty}\frac{1}{p_i^{\frac{j}{m}}}=\prod_{i=1}^{t}\frac{1}{1-p_i^{-\frac{1}{m}}}.$$
又因为存在正整数 N, 使得当 $n\ge N$ 时, 有 $f(n)^{\frac{1}{m}}<2b_m^{\frac{1}{m}}n$, 因此
$$\sum_{n=N}^{\infty}\frac{1}{f(n)^{\frac{1}{m}}}>\frac{1}{2b_m^{\frac{1}{m}}}\sum_{n=N}^{\infty}\frac{1}{n},$$
不存在上界, 这就导致了矛盾. $\square$

评注 本题的结论是定理 1 的推广, 当 n 充分大时, 显然 $f(n)$ 是单调的.

正如定理 1 可以看作 Euclid 的手法的再次成功, 本题的结论, 也可以视作 Euler 对于素数无限的证明的推广. 在这种有分析色彩的数论问题中, 取倒数相加或许是一种有力的手段. Dirichlet 在证明他的著名定理——算术级数中有无穷多素数时, 采用的正是这一方法.

在素数无限的证明中, 下面的这种证法称不上简单, 也并不算很漂亮. 假设只有有限个素数 $p_1,\ldots,p_k$, 则 $1,2,\ldots,N$ 中每个数可以写成 $\prod\limits_{i=1}^{k}p_i^{\alpha_i}$ 的形式, 且
$$2^{\alpha_1+\cdots+\alpha_k}\le\prod_{i=1}^{k}p_i^{\alpha_i}\le N,$$
因此 $\alpha_i\le\alpha_1+\cdots+\alpha_k\le\log_2 N$. 于是 $\prod\limits_{i=1}^{k}p_i^{\alpha_i}$ 最多能表示出 $(\log_2 N+1)^k$ 个不同的数, 因此 $(\log_2 N+1)^k\ge N$, 取 N 充分大, 这不可能成立.

但这种证法有趣之处在于: 它表明, 如果正整数列中只有有限个素数, 那么它的增长速度似乎太慢了一些. 认真追究这个证法成功的原因, 或许我们不难得到如下的定理.

定理 3 (Christian Elsholtz[2]). 令 $S=(s_1,s_2,s_3,\ldots)$ 为一整数序列. 称

(1) S 是几乎单射的, 如果存在常数 c, 每个值至多出现 c 次;

(2) S 是次指数增长的, 如果存在一个函数 $f:\mathbb{N}\to\mathbb{R}^+$, 满足 $\lim\limits_{n\to\infty}\frac{f(n)}{\log_2 n}=0$, 使得 $|s_n|\le 2^{2^{f(n)}}$ 对所有 n 成立.

设整数数列 $\{a_n\}$ 是几乎单射的, 且是次指数增长的, 则 $\{a_n\}$ 中有无穷多个素因子.

证明 设 $f(n)$ 都是单调不减的, 否则用 $F(n)=\max\limits_{i\le n}f(i)$ 代替 $f(n)$ 不影响问题. 假设 $\{a_n\}$ 只有有限个素因数 $p_1,\ldots,p_k$. 对任意的 n, 设 $a_n=\varepsilon_n\prod\limits_{i=1}^{k}p_i^{\alpha_i}$, 其中 $\varepsilon_n\in\{-1,0,1\},\alpha_i\ge 0$, 则有
$$2^{\alpha_1+\cdots+\alpha_k}\le a_n\le 2^{2^{f(n)}}.$$

以 2 为底取对数, 我们得到

$$0 \le \alpha_i \le \alpha_1 + \cdots + \alpha_k \le 2^{f(n)},\ 1 \le i \le k.$$

因此, 每个 $\alpha_i = \alpha_i(n)$ 有不超过 $2^{f(n)} + 1$ 个不同的可能值, 而 f 是单调的. 这表明, 对于给定的 N, 在 $|a_1|, |a_2|, \ldots, |a_N|$ 中最多有 $(2^{f(N)} + 1)^k$ 个不同的数.

另一方面, 由于 $\{a_n\}$ 是几乎单射的, 序列中只有 c 项可以为 0, 每个非零的绝对值至多可以出现 $2c$ 次. 于是, 在 $|a_1|, |a_2|, \ldots, |a_N|$ 中至少有 $\frac{N-c}{2c}$ 个不同的数. 放在一起, 我们得到 $\frac{N-c}{2c} \le (2^{f(N)+1})^k$, 由于 $\frac{f(n)}{\log_2 n} \to 0$, 对于充分大的 N, 这不可能成立, 矛盾. □

评注 显然, 这个定理蕴含了定理 1 和定理 2.

这里的次指数增长, 实际上比指数增长要慢很多. 它并不能理解为存在一个函数 $g: \mathbb{N} \to \mathbb{R}^+$, 满足 $\lim\limits_{n\to\infty} \frac{g(n)}{n} = 0$, 使得 $|s_n| \le 2^{g(n)}$ 对 n 所有成立. 比如, 取 $g(n) = \sqrt{n}$, 那么 $\lim\limits_{n\to\infty} \frac{g(n)}{n} = 0$, 但是 $\lim\limits_{n\to\infty} \frac{\log_2 g(n)}{\log_2 n} = \frac{1}{2}$, 然而 $\lim\limits_{n\to\infty} \frac{f(n)}{\log_2 n} = 0$ 不能改进了. 对前 T 个素数 $p_1, \ldots, p_T$, 考虑所有形如 $\prod\limits_{i=1}^{T} p_i^{\alpha_i}$ 的数按从小到大顺序排列得到的数列, 它的增长速度大概是 $2^{2^{f(n)}}$, 其中 $\frac{f(n)}{\log_2 n} \approx \frac{1}{T}$, 但这样的数列只有有限个素因子.

下面这道最近的赛题, 应用上面的定理可以很快做出.

题 1 (2016 USA TST). *求所有大于 1 的整数 C, 使得存在由两两不同的正整数构成的无穷数列 $a_1, \ldots, a_k$, 对于任意正整数 k, 都有 $a_{k+1}^k \mid C^k a_1 a_2 \ldots a_k$.*

解 我们来证明 $\{a_n\}$ 是次指数增长的, 由定理 3 即得不存在这样的 C.

令 $b_n = \log_2 a_n$, $A = \log_2 C$, 有 $kb_{k+1} \le kA + b_1 + \cdots + b_k$. 令 $A_n = \frac{b_1 + \cdots + b_n}{n}$, 有 $A_{k+1} \le A_k + \frac{A}{k+1}$, 因此存在一个正实数 M, 使得

$$A_n \le M \log_2 n,\ b_n \le A + A_{n-1} \le A + M \log_2 n,$$

因此 $\lim\limits_{n\to\infty} \frac{\log_2 b_n}{\log_2 n} = 0$. □

评注 事实上 b_n 是对数级增长的, 于是 a_n 是幂级增长的, 应用定理 2 也可以解决本题.

可以看出, 研究素数的幂起到了非常重要的作用. 事实上, 当我们假设数列 $\{a_n\}$ 中仅有有限个素数 $p_1, \ldots, p_k$ 时, 使用这个反证假设, 得到的唯一信息就是, 每一个 a_n 都可以写成 $\prod\limits_{i=1}^{k} p_i^{\alpha_i}$ 的形式. 那么自然就会去考虑某个素数幂次的大小.

下面的题便是一个利用阶乘性质去分析幂次的例子.

题 2 (2012 RMO[3]). 记 $S_n = 1! + 2! + \cdots + n!$, 证明: 存在 $n \in \mathbb{N}$, 使得 S_n 有大于 10^{2012} 的素因数.

证明 对任意素数 p, 任意正整数 n, 用 $v_p(n)$ 表示 n 的质因数分解中素数 p 的幂次.

注意到, 若 $v_p(n) \neq v_p(k)$, 则 $v_p(n \pm k) = \min\{v_p(n), v_p(k)\}$. 由此得到下面的引理:

引理 如果存在某个正整数 n 满足 $v_p(S_n) < v_p((n+1)!)$, 则对于任意的 $k \geq n$, 有 $v_p(S_k) = v_p(S_n)$.

令 $P = 10^{2012}$. 假设对于任意正整数 n, S_n 的所有素因子都小于 P. 对于任意素数 $p < P$, 如果存在某个正整数 m, 使得 $v_p(S_m) < v_p((m+1)!)$. 根据引理, 存在正整数 a_p, 使得对任意的正整数 n, 有 $v_p(S_n) \leq a_p$. 我们称这样的素数为 "小素数", 所有小于 P 且不是小素数的素数称为 "大素数".

取定正整数 M, 使得不等式 $M > p^{a_p}$ 对任意小素数成立.

对任意一个大素数 p, 如果 $n+2$ 是 p 的倍数, 则由引理,

$$v_p(S_{n+1}) \geq v_p((n+2)!) > v_p((n+1)!),$$

这推出 (注意到 2 显然是小素数)

$$v_p(S_n) = v_p(S_{n+1} - (n+1)!) = v_p((n+1)!) = v_p(n!).$$

令 $N = MP! - 2$, 则上述论证表明 $v_p(S_N) = v_p(N!)$ 对任意大素数 p 成立. 又因为 $N \geq M$, $v_p(S_N) \leq v_p(p^{a_p}) \leq v_p(N!)$, 而 S_N 所有的素因数不是小素数就是大素数, 这表明 $S_N \leq N!$, 矛盾. □

评注 本题还有别的一些证明, 但大都用到了公式: $v_p(n!) = \sum\limits_{k=1}^{\infty} [\frac{n}{p^k}]$.

而上面的证明, 也即标准答案中的做法, 用到的知识是非常少的. 它几乎只利用了如下的事实:

若 $v_p(a) \neq v_p(b)$, 则 $v_p(a+b) = \min\{v_p(a), v_p(b)\}$.

然后围绕阶乘的性质分析其中素数的幂次, 最终得到矛盾.

在分析幂次时, 一种常见的手法是, 考虑数列中的若干个数, 其中必有两个数, 素数部分最大者对应着同一个素数, 这样, 两个数的最大公约数就大于较小数的 k 次根. 下题便是这样一个例子.

题 3 (2009 Iran TST). 设 a 是一个给定的正整数, 证明: 集合 $S = \{2^{2^n} + a \mid n \in \mathbb{N}^+\}$ 中的整数有无穷多个不同的素因子.

证明 设 S 中全体素因子为 $p_1 < \ldots < p_k$. 对任意的 $r \in \mathbb{N}^+$, 取充分大的正整数 N, 使得 $2^{2^N} + a > (\prod\limits_{i=1}^{k} p_i)^r$. 令

$$B = \{2^{2^N} + a, 2^{2^{N+1}} + a, \cdots, 2^{2^{N+k}} + a\}.$$

对任意 $b \in B$, 有一个素因子的幂次大于 r, 而 $|B| = k+1$, 因此存在 u, v, 使得对于某个 p_i, 有 $p_i^r \mid 2^{2^u} + a$, $p_i^r \mid 2^{2^{u+v}} + a$. 而

$$2^{2^u} + a \mid (2^{2^u})^{2^v} - a^{2^v} = 2^{2^{u+v}} - a^{2^v},$$

因此 $p_i^r \mid a^{2^v} + a$, 于是 $p_i^r \leq a^{2^v} + a \leq a^{2^k} + a$, 但 k 是常数, 矛盾. □

评注 我们可以走得更远一些. 注意到 $\{2^{2^n} + a\}$ 可以看做由 $\{2^{2^n}\}$ 平移得到的, 而后者仅存在有限个素因子. 下面, 我们以一个强大的定理作结此文.

定理 4 (Kobayashi). 若无界正整数列 $\{a_n\}_{n\geq 0}$ 只存在有限个素因子, 则对 $a \in \mathbb{Z},\ a \neq 0$, 数列 $\{a_n - a\}_{n\geq 0}$ 有无穷多个素因子.

证明 我们不加证明的使用 Thue 定理[4]:

引理 (Thue) 设 $n \geq 3$, $f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ 是一个整系数的 n 次(有理数域上) 既约多项式, 则不定方程

$$H(x, y) = a_n x^n + a_{n-1} x^{n-1} y + \cdots + a_1 x y^{n-1} + a_0 y^n = C$$

仅有有限多组整数解 x, y, 其中 C 是给定的整数.

回到我们的问题. 设 $\{a_n\}$ 的素因子仅有 $p_1, \ldots, p_r$, 且 $\{a_n - a\}$ 的素因子仅有 $q_1, q_2, \ldots, q_s$. 由于 $\{a_n\}$ 无界, 因此 $\prod\limits_{i=1}^{r} p_i^{\alpha_i} - \prod\limits_{j=1}^{s} q_j^{\beta_j}$ 有无穷多组非负整数解.

考虑向量 $v = (\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_s)$, 将它们模 3 分类, 必有一类有无穷个元素. 这样的一类解, 可以写成 $Ax^3 - By^3 = a$ 的形式. 其中 A, B 是确定的正整数, 并且 A, B 中每个素因子的幂次都不超过 2.

若 $A = B$, 有 $x^2 + xy + y^2 \mid a$, 因此 $x^2 + xy + y^2 \leq a$, 只能有有限组解. 若 $A \neq B$, 则 $\sqrt[3]{\frac{A}{B}} \notin \mathbb{Q}$, 因此 $Ax^3 - By^3$ 既约, 根据 Thue 定理, 只能有有限组解. 这就导致了矛盾. 故 $\{a_n - a\}_{n\geq 0}$ 有无穷多个素因子. □

参考文献

[1] 余红兵. 奥数教程 · 高三年级 [M]. 第六版, 上海: 华东师范大学出版社, 2014.4.

[2] Martin Aigner, Gunter M. Ziegler. 数学天书中的证明 [M]. 第五版, 北京: 高

等教育出版社, 2016. 3.

[3] 2012 年 IMO 中国国家集训队教练组. 走向 IMO: 数学奥林匹克试题集锦 (2012) [M]. 上海: 华东师范大学出版社, 2012. 8.

[4] 柯召, 孙琦. 数论讲义 [M]. 第 2 版, 北京: 高等教育出版社, 2003. 5.