Riemann 猜想漫谈

卢昌海

If you could be the Devil and offer a mathematician to sell his soul for the proof of one theorem ——what theorem would most mathematicians ask for? I think it would be the Riemann Hypothesis.

H. Montgomery

一. **Hardy 的电报**

让我们从一则小故事开始我们的Riemann 猜想之旅吧。故事发生在大约七十年前，当时英国有一位很著名的数学家叫做 Godfrey Hardy (1877-1947)，他是两百年来英国数学界的一位“勇者”。为什么说他是勇者呢？因为在十七世纪的时候，英国的数学家与欧洲大陆的数学家之间发生了一场剧烈的论战。论战的话题是谁先发明了微积分。论战的当事人一边是英国的科学泰斗 Isaac Newton(1642-1727)， 另一边是欧洲大陆(德国)的哲学及数学家 Gottfried Leibniz (1646-1716)。这一场论战打下来，两边筋疲力尽自不待言，还大伤了和气，留下了旷日持久的后遗症。英国的许多数学家开始排斥起来自欧洲大陆的数学进展。一场争论演变到这样的一个地步，英国数学界的集体荣誉及尊严、Newton 的赫赫威名便都成了负资产，英国的数学在保守的舞步中走起了下坡路。

这下坡路一走便是两百年。

在这样的一个背景下，在复数理论还被一些英国数学家视为来自欧洲大陆的危险概念的时候，土生土长的英国数学家 Hardy 却对来自欧洲大陆(德国——又是德国)、有着复变函数色彩的数学猜想——Riemann 猜想——产生了浓厚的兴趣，积极地研究它，并且取得了令欧洲大陆数学界为之震动的成就 (这一成就将在后文中介绍)，算得上是勇者所为。

当时 Hardy 在丹麦有一位很好的数学家朋友叫做 Harald Bohr (1887-1951)， 他是著名量子物理学家 Niels Bohr 的弟弟。Bohr 对 Riemann 猜想也有浓厚的兴趣，曾与德国数学家 Edmund Landau (1877-1938) 一起研究 Riemann 猜想 (他们的研究成果也将在后文中介绍)。Hardy 很喜欢与 Bohr 共度暑假，一起讨论 Riemann 猜想，常常待到假期将尽才匆匆赶回英国。结果有一次当他赶到码头时，发现只剩下一条小船可以乘坐了。在汪洋大海中乘坐一条小船可不是闹着玩的事情，弄得好算是浪漫刺激，弄不好就得葬身鱼腹。信奉上帝的乘客们此时都忙着祈求上帝的保佑。Hardy 却是一个坚决不信上帝的人，不仅不信上帝，有一年还把向大众证明上帝不存在列入自己的年度六大心愿之中，且排名第三 (排名第一的是证明 Riemann 猜想)。不过在这生死攸关的时刻 Hardy 也没闲着，他给 Bohr 发去了一封电报，电报上只有一句话：

“我已经证明了 Riemann 猜想！”

Hardy 为什么要发这么一个电报呢？回到英国后他向 Bohr 解释了原因，他说如果那次他乘坐的船真的沉没了，那人们就只好相信他真的证明了 Riemann 猜想，但他知道上帝是肯定不会把这么巨大的荣誉送给他——一个坚决不信上帝的人——的，因此上帝一定不会让他的小船沉没的。[注一]

上帝果然没有舍得让 Hardy 的小船沉没。自那以后又过去了七十来个年头，吝啬的上帝仍然没有物色到一个可以承受这么大荣誉的人。

二. **Riemann ζ 函数与 Riemann 猜想**

那么这个让上帝如此吝啬的Riemann 猜想究竟是一个什么样的猜想呢？在回答这个问题之前我们先来介绍一个函数： Riemann ζ 函数。这个函数虽然挂着 Riemann 的大名，

却不是 Riemann 提出的。但是 Riemann 虽然不是这一函数的提出者，他的工作却大大加深了人们对这一函数的理解，为其在数学与物理上的广泛运用奠定了基础。后人为了纪念 Riemann 的卓越贡献，就用他的名字命名了这一函数。[注二]

Riemann ζ 函数 ζ(s) 是级数表达式 (n 为自然数)

$$\zeta(s) = \sum_n n^{-s} (\mathrm{Re}(s) > 1)$$

在复平面上的解析延拓。之所以需要解析延拓，是因为上面这一表达式——如我们已经注明的——只适用于复平面上 Re(s) > 1 的区域 (否则级数不收敛)。Riemann 找到了上面这一表达式的解析延拓 (当然 Riemann 没有使用 "解析延拓" 这一现代复变函数论的术语)。运用路径积分，解析延拓后的 Riemann ζ 函数可以表示为：

$$\zeta(s) = \frac{\Gamma(1-s)}{2\pi i} \int_{\infty}^{\infty} \frac{(-z)^s}{e^z - 1} \frac{dz}{z}$$

式中的积分环绕正实轴进行(即从∞出发，沿实轴上方积分至原点附近，环绕原点积分至实轴下方，再沿实轴下方积分至∞——离实轴的距离及环绕原点的半径均趋于 0)；式中的 Γ 函数 Γ(s) 是阶乘函数在复平面上的推广，对于正整数 s>1：Γ(s)=(s-1)!。可以证明，这一积分表达式除了在 s=1 处有一个简单极点外在整个复平面上解析。这就是 Riemann ζ 函数的完整定义。

运用上面的积分表达式可以证明，Riemann ζ 函数满足以下关系式：

$$\zeta(s) = 2\Gamma(1-s)(2\pi)^{s-1}\sin(\pi s/2)\zeta(1-s)$$

从这个关系式中不难发现，Riemann ζ 函数在 s=-2n (n 为自然数)取值为零——因为 sin(πs/2)为零[注三]。复平面上的这种使 Riemann ζ 函数取值为零的点被称为 Riemann ζ 函数的零点。因此 s=-2n (n 为自然数) 是 Riemann ζ 函数的零点。这些分布有序的零点性质十分简单，被称为 Riemann ζ 函数的平凡零点 (trivial zeros)。除了这些平凡零点外， Riemann ζ 函数还有许多其它的零点，那些零点被称为非平凡零点。 对 Riemann ζ 函数非平凡零点的研究构成了现代数学中最艰深的课题之一。我们所要讨论的 Riemann 猜想就是关于这些非平凡零点的猜想，在这里我们先把它的内容表述一下， 然后再叙述它的来笼去脉：

Riemann 猜想：Riemann ζ 函数的所有非平凡零点 (non-trivial zeros) 都位于复平面上 Re(s)=1/2 的直线上。

在 Riemann 猜想的研究中数学家们把复平面上 Re(s)=1/2 的直线称为 critical line，运用这一术语，Riemann 猜想也可以表述为：Riemann ζ 函数的所有非平凡零点都位于 critical line 上。

这就是 Riemann 猜想的内容，它是 Riemann 在 1859 年提出的。从其表述上看，Riemann 猜想似乎是一个有关复变函数的命题，但我们很快将会看到，它其实却是一曲有关素数分布的神秘乐章。

注释

[注一] 这个故事让我想起一句有趣的无神论者的祈祷语：God, if there is one, save my soul if I have one (上帝啊，如果你存在的话，拯救我的灵魂吧，如果我有灵魂的话)。

[注二] 远在 Riemann 之前，Riemann ζ 函数（当然那时还不叫这个名字）的级数表达式就已经出现在了数学文献中，但是那些表达式中函数的定义域较小。Riemann 把 Riemann ζ 函数的定义域大大地延拓了，这一点对于 Riemann 猜想的表述及研究具有重要的意义。仅凭这一点，即便把 Riemann 称为 Riemann ζ 函数的提出者之一，也并不过份。

[注三] sin(πs/2) 在 s=0 及 s=2n (n 为自然数) 时也为零，但是 s=0 时 ζ(1-s) 有极点，s=2n (n 为自然数) 时 Γ(1-s) 有极点，因此只有在 s=-2n (n 为自然数) 时可以由 sin(πs/2)=0 推知 Riemann ζ 函数的取值为零。

三. 素数的分布

一个复数域上的函数——Riemann ζ 函数——的非平凡零点（以后将简称为零点）的分布怎么会与风马牛不相及的自然数域中的素数分布产生关联呢？这还得从 Euler 乘积公式谈起。

我们知道，早在古希腊时代，Euclid 就用精彩的反证法证明了素数有无穷多个。随着数论研究的深入，人们很自然地对这些素数在自然数域中的分布产生了越来越浓厚的兴趣。1737 年，著名数学家 Leonhard Euler(1707-1783)在圣彼得堡科学院(St. Petersburg Academy) 发表了一个极为重要的公式，为数学家们研究素数分布的规律奠定了基础。这个公式就是 Euler 乘积公式：

$$\sum_n n^{-s} = \prod_p (1 - p^{-s})^{-1}$$

公式中左边的求和对所有的自然数进行，右边的连乘积对所有的素数进行。可以证明，这个公式对所有 Re(s)>1 的复数 s 都成立。这个公式的左边正是我们在上文中介绍过的 Riemann ζ 函数，而右边则是一个纯粹有关素数（且包含所有素数）的表达式，这样的形式正是 Riemann ζ 函数与素数分布之间存在关联的征兆。那么这个公式究竟蕴涵着有关素数分布的什么样的信息呢？Riemann ζ 函数的零点又是如何出现在这种关联之中的呢？ 这就是本节及下节所要介绍的内容。

Euler 本人率先对这个公式所蕴涵的信息进行了研究。他注意到在 s=1 的时候，公式的左边 $\sum_n n^{-1}$ 是一个发散级数（这是一个著名的发散级数，称为调和级数），这个级数以对数方式发散。这些对于 Euler 来说都是不陌生的。为了处理公式右边的连乘积，他对公式两边同时取了对数，于是连乘积变成了求和， 由此他得到：

$$\ln(\sum_n n^{-1}) = -\sum_p \ln(1 - p^{-1}) = \sum_p \sum_{k=1}^{\infty} \frac{p^{-k}}{k}$$

由于上式右端括号中除第一项外所有其它各项的求和都收敛，而且这些求和的结果累加在一起仍然收敛（有兴趣的读者不妨自己证明一下）。因此右边只有第一项的求和是发散的。由此 Euler 得到了这样一个有趣的渐近表达式：

$$\sum_{p \le x} p^{-1} \sim \ln\ln x$$

这个结果——即 $\sum_{p \le x} p^{-1}$ 以 lnlnx 的方式发散——是继 Euclid 证明素数有无穷多个以来有关素数的又一个重要的研究结果。它同时也是对素数有无穷多个这一命题的一种崭新的证明（因为假如素数只有有限多个，则求和就只有有限多项，不可能发散）。但 Euler 的这一新证明所包含的内容要远远多于 Euclid 的证明，因为它表明素数不仅有无穷多个，而且其

分布要比许多同样也是无穷的序列——比如n^2 序列——密集得多(因为后者的倒数之和收敛)。不仅如此，如果我们进一步注意到上式的右端可以改写为一个积分表达式：

$$\ln\ln x \sim \int \frac{1}{x\ln x}dx$$

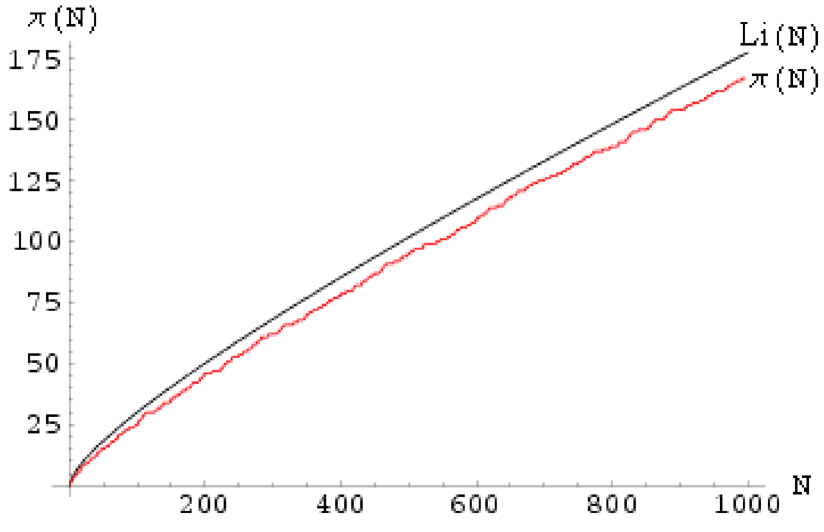
而左端通过引进一个素数分布的密度函数 ρ(x)——它给出在 x 附近单位区间内发现素数的几率——也可以改写为一个积分表达式：

$$\sum_{p\le x} p^{-1} \sim \int \frac{\rho(x)}{x}dx$$

将这两个积分表达式进行比较，不难猜测到素数的分布密度为 $\rho(x)\sim 1/\ln(x)$， 从而在 x 以内的素数个数——通常用 $\pi(x)$ 表示为：

$$\pi(x) \sim \mathrm{Li}(x)$$

其中 $\mathrm{Li}(x) \equiv \int \frac{1}{\ln x} dx$ 是对数积分函数[注一]。这正是著名的素数定理 (当然这种粗略的推理并不构成素数定理的证明)。因此 Euler 发现的这个结果可以说是一扇通向素数定理的暗门。可惜 Euler 本人并没有沿着上面的思路走，从而错过了这扇暗门，数学家们提出素数定理的时间也因此而延后了几十年。



素数分布与素数定理

提出素数定理的这份荣誉最终落到了另外两位数学家的肩上：他们是德国数学家 Friedrich Gauss (1777-1855)和法国数学家 Adrien-Marie Legendre (1752-1833)。

Gauss 对素数分布的研究始于 1792 到 1793 年间,那时他才 15 岁。在那期间，每当“无所事事”的时候 Gauss 就会挑上几个长度为一千的自然数区间，计算这些区间中的素数个数，并进行比较。在做过了大量的计算和比较后，Gauss 发现素数分布的密度可以近似地用对数函数的倒数来描述，即 $\rho(x)\sim 1/\ln x$，这正是上面提到的素数定理的主要内容。但是 Gauss 并没有发表这一结果。Gauss 是一个追求完美的数学家，他很少发表自己认为还不够完美的结果，而他的数学思想和灵感犹如浩瀚奔腾的江水，汹涌激荡，常常让他还没来得及将一个研究结果完美化就又展开了新课题的研究。因此 Gauss 一生所做的数学研

究远远多过他正式发表的。但是另一方面，Gauss 常常会用其它的方式——比如通过书信——透露自己的某些未发表的研究成果，他的这一做法给一些与他同时代的数学家带来了不小的尴尬。其中 “受灾” 较为深重的一位便是 Legendre。这位法国数学家在 1806 年率先发表了线性拟合中的最小平方法，不料 Gauss 在 1809 出版的一部著作中提到自己曾在 1794 年 (即比 Legendre 早了 12 年) 就发现了同样的方法。使 Legendre 极为不快。

有道是：不是冤家不聚首。在素数定理的提出上，可怜的 Legendre 又一次不幸地与数学巨匠 Gauss 撞到了一起。 Legendre 在 1798 年发表了自己关于素数分布的研究，这是数学史上有关素数定理的最早的文献[注二]。由于 Gauss 没有发表自己的研究结果，Legendre 便理所当然地成为了素数定理的提出者。 Legendre 的这个优先权一共维持了 51 年。到了 1849 年 Gauss 在给德国天文学家 Johann Encke (1791-1865) 的一封信中提到了自己在 1792 至 1793 年间的研究，从而把尘封了半个世纪的优先权从 Legendre 的口袋中勾了出来，挂到了自己已经鼓鼓囊囊的腰包上。

幸运的是，Gauss 给 Encke 写信的时候 Legendre 已经去世十六年了，他用最无奈的方法避免了再次遭受残酷的打击。

无论 Gauss 还是 Legendre，他们对于素数分布规律的研究都是以猜测的形式提出的 (Legendre 的研究带有一定的推理成份,但离证明仍相距甚远)。 因此确切地说，素数定理在那时只是一个猜想——素数猜想， 我们所说的提出素数定理指的也只是提出素数猜想。素数定理的数学证明直到一个世纪之后的 1896 年，才由法国数学家 Jacques Hadamard (1865-1963) 与比利时数学家 Charles de la Vallée-Poussin (1866-1962) 彼此独立地给出。他们的证明与 Riemann 猜想有着很深的渊源，其中 Hadamard 的证明出现的时机和场合还富有很大的戏剧性，这些我们将在后文中加以叙述。

素数定理是简洁而且优美的，但是它对于素数分布的描述仍然是比较粗略的，它给出的只是素数分布的一个渐近形式——也就是说是当 x 趋于无穷时的分布形式。从前面有关素数分布与素数定理的图示中我们也可以看到，$\pi(x)$ 与 Li(x) 之间是有偏差的，而且这种偏差的绝对值随着 x 的增加似有持续增加的趋势 (所幸的是，这种偏差的增加与 $\pi(x)$ 及 Li(x) 本身的增加相比仍然是微不足道的——否则素数定理也就不成立了)[注三]。

那么有没有一个公式可以比素数定理更精确地描述素数的分布呢？这便是 Riemann 在 1859 年想要回答的问题。那一年是 Gauss 去世后的第五年，32 岁的 Riemann 继 Johann Dirichlet (1805-1859) 之后成为了 Gauss 在 Göttingen 大学的继任者。 同年八月十一日，他被选为柏林科学院 (Berlin Academy) 的通信院士 (Corresponding Member)。 作为对这一崇高荣誉的回报，Riemann 向柏林科学院提交了一篇论文。这是一篇只有短短八页的论文， 标题是： 论小于给定数值的素数个数。 正是这篇论文将 Euler 乘积公式蕴涵的信息破译得淋漓尽致，也正是这篇论文将 Riemann ζ 函数的零点分布与素数的分布联系在了一起。

这篇论文注定要把人们对素数分布的研究推向壮丽的巅峰，并为后世的数学家们留下一个魅力无穷的伟大谜团。

注释

[注一] 对数积分函数 Li(x) 的确切定义是 1/lnx 在 0 到 x 之间定积分的 Cauchy 主值。对于素数定理来说，人们关心的是 Li(x) 在 $x\to\infty$ 时的渐近行为，这时候积分的下限并不重要，因此人们在素数定理的研究中有时把 Li(x) 的积分下限取为 2 而不是 0，这样可以使被积函数在积分区间内没有奇点。

[注二] Legendre 提出的素数定理采用的是代数表达式： $\pi(x) \sim x/(\ln x-1.08366)$， 它

与积分形式的素数定理在渐近意义上是等价的。

[注三] 从图上以及从更大范围的计算中人们发现 Li(x)-π(x) 总是大于零，以致于有人猜测 Li(x) 不仅是素数分布的渐近形式，而且还是其严格上界。这种猜测在 1904 年被英国数学家 John Littlewood (1885-1977) 所推翻。Littlewood 证明了 Li(x)-π(x) 是一个在正与负之间震荡无穷多次的函数。

四. Riemann 的论文——基本思路

终于到了 Riemann 的论文登场的时候！如果让数学家们来评选几篇数学史上意义深远却又最为难读的论文，那么我想 Riemann 1859 年的那篇 "论小于给定数值的素数个数" 就算不名列榜首，起码也要挤身三甲。现在就让我们来一起领略一下那篇数学史上出名难啃的论文的主要内容。我们的叙述将采用较为现代的术语和方式，所用的记号将与前文保持一致，因此与 Riemann 的原始论文不尽相同 (但主要思路是一致的)。这一点请有兴趣阅读 Riemann 原文的读者注意。

如上节所述，Euler 乘积公式：

$$\zeta(s) \equiv \sum_n n^{-s} = \prod_p (1 - p^{-s})^{-1}$$

是研究素数分布规律的基础。Riemann 的研究也以这一公式作为起点。为了消除右边的连乘积，Euler 曾对公式两边取对数，Riemann 也如法泡制 (看来连乘积真是人见人恨)，从而得到：

$$\ln\zeta(s) = \ln(\sum_n n^{-1}) = -\sum_p \ln(1 - p^{-1}) = \sum_p \sum_{k=1}^{\infty} \frac{p^{-k}}{k}$$

过了这步，两人就分道扬镳了：Euler 如我们在上节所见——小试身手，证明了素数有无穷多个，然后就喜滋滋地鸣金收兵了；而 Riemann 则沿着一条布满荆棘的道路继续走了下去，走出了素数研究的一片崭新的天地。

可以证明，上式右边的双重求和在复平面上 Re(s)>1 的区域内是绝对收敛的，并且可以改写成 Stieltjes 积分(有兴趣的读者可自行证明)：

$$\ln\zeta(s) = \int_0^{\infty} x^{-s}\, \mathrm{d}J(x)$$

其中 J(x) 是一个特殊的阶梯函数，它在 x=0 取值为零，以后每越过一个素数就增加 1，每越过一个素数的平方就增加 1/2，……，每越过一个素数的 n 次方就增加 1/n，……。在 J(x) 不连续的点 (即 x 等于素数、素数的平方、……、素数的 n 次方……的点)上其函数值用 $J(x) = \frac{J(x^-) + J(x^+)}{2}$ 来定义。显然，这样的一个阶梯函数可以用素数分布函数 $\pi(x)$ 表示为：

$$J(x) = \sum_n \frac{\pi(x^{1/n})}{n}$$

对上述 Stieltjes 积分进行一次分部积分便可得到：

$$\ln\zeta(s) = s\int_0^\infty J(x)x^{-s-1}\mathrm{d}x$$

这个公式的左边是 Riemann ζ 函数的自然对数，右边则是对 J(x)——一个与素数分布函数 π(x) 有直接关系的函数——的积分，它可以被视为 Euler 乘积公式的积分形式。我们得到这一结果的方法与 Riemann 有所不同，Riemann 发表论文时还没有 Stieltjes 积分——那时候 Thomas Stieltjes (1856-1894) 才三岁。

如果说传统形式下的 Euler 乘积公式只是 Riemann ζ 函数与素数分布之间存在关联的征兆，那么在这个积分形式的 Euler 乘积公式下这两者之间的关联就已是确凿无疑并且完全定量了。接下来首先要做的显然是从上述积分中解出 J(x) 来，这在当时的数学背景下并不容易，但却难不倒象 Riemann 这样的复变函数论大师。他解出的 J(x) 是 (学过复变函数论的读者不妨试着证明一下)：

$$J(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{\ln\zeta(z)}{z}x^z\mathrm{d}z$$

其中 a 为大于 1 的实数。这是一个条件收敛的积分，它的确切定义是从 a-ib 积分到 a+ib (b 为正实数)，然后取 b→∞的极限。当 Riemann 写下这个公式时，只是轻描淡写地提了一句：这是完全普遍的。听上去就象是在叙述一个尽人皆知的简单事实。而事实上，与 Riemann 所说的普遍性相匹配的完整结果直到四十年后才由芬兰数学家 Robert Mellin (1854-1933) 所发表，现在被称为 Mellin 变换(Mellin Transform)。象这样一种被 Riemann 随手写下、却让数学界花费几十年甚至上百年的时间才能证明的命题在 Riemann 的那篇论文中还有好几处。这是 Riemann 那篇论文的一个极为突出的特点：它有一种高屋建瓴的宏伟视野，远远地超越了同时代的其他数学文献。它那高度浓缩的文句背后包含着的极为丰富的数学结果，让后世的数学家们陷入了漫长的深思之中。直到今天，我们的数学在整体上虽已远非 Riemann 时代可比，但数学家们仍未能完全理解 Riemann 在那篇短短八页的简短论文中显露出的全部智慧。J(x)的表达式是我们碰到的 Riemann 论文中的结果超前于时代的第一个例子[注一]，在下一节中我们将遇到其它例子。

在一代代的后世数学家们为那些被 Riemann 省略掉的证明而失眠的时候，他们中的一些也许会联想到 Pierre de Fermat (1601-1665)。这位法国数学家在 Diophantus 的«Arithmetica» 页边上写下著名的 Fermat 猜想 (Fermat's Last Theorem) 的时候，随手加了一句话："我发现了一个真正出色的证明，可惜页边太窄写不下来" [注二]。令人尴尬的是，Fermat 的猜想 1670 年被他儿子公诸于世(那时他本人已经去世) 以来，竟然难倒整个数学界长达 324 年之久，直到 1994 年才被英国数学家 Andrew Wiles 所证明。但 Wiles 的证明篇幅浩繁，莫说在 «Arithmetica» 的页边上写不下来，即便把整个大英百科全书的页边加起来，也未必写得下来。现在人们普遍认为，Fermat 并没有找到 Fermat 猜想的证明，他自以为找到的那个"真正出色的证明" 只是三百多年间无数个错误证明中的一个[注三]。那么 Riemann 的情形会不会也象 Fermat 一样呢？他的那些省略掉的证明会不会也象 Fermat 的那个"真正出色的证明" 一样呢？从目前人们对 Riemann 的研究来看，答案是否定的。Riemann 作为堪与 Gauss 齐名的有史以来最伟大的数学家之一，他的水平远非 Fermat 可比。而且人们在对 Riemann 的部分手稿进行研究时发现，Riemann 对自己论文中的许多语焉不详的命题是做过扎实的演算和证明的，只不过他和 Gauss 一样追求完美，发表的东西远远少于自己研究过的。更令人钦佩的是，Riemann 手稿中一些演算和证明哪怕是时隔了几十年之后才被整理出来，却仍然大大超越当时数学界的水平。因此我们有一定的理由相信，Riemann 在论文中以陈述而不是猜测的语气表述的内容——不论有没有给出证明——都是有着深入的演算和证明背景的。

好了，现在回到 J(x) 的表达式来，这个表达式给出了 J(x) 与 Riemann ζ 函数之间的确切关联。换句话说，只要知道了 ζ(s)，通过这个表达式原则上就可以计算出 J(x)。知道了 J(x)，下一步显然就是计算 $\pi(x)$。这并不困难， 因为上面提到的 J(x) 与 $\pi(x)$ 之间的关系式可以通过所谓的 Möbius 反演(Möbius Inversion) 解出，结果为：

$$\pi(x)=\sum_n \frac{\mu(n)}{n} J(n^{1/n})$$

其中 μ(n) 被称为 Möbius 函数， 它的取值如下：

μ(1)= 1

$\mu(n)=(-1)^l$ 如果 n 是 l 个不同素数的乘积。

μ(n)=0　　其他

因此知道 J(x) 就可以计算出 $\pi(x)$，即素数的分布函数。把这些步骤连接在一起，我们看到，从 ζ(x) 到 J(x)，再从 J(x)到 $\pi(x)$，素数分布的秘密完全定量地蕴涵在了 Riemann ζ 函数之中。这就是 Riemann 研究素数分布的基本思路。在下一节中，我们将进一步深入 Riemann 的论文，让那些千呼万唤犹未露面的 Riemann ζ 函数的零点显露在我们的镁光灯下。

二零零三年十二月六日写于纽约

注释

[注一]为了先把 Riemann 论文的思路表述清楚，我们对叙述的顺序作了调整，因此这里所说的“第一个例子”是相对于我们的叙述而言的。 在 Riemann 的原始论文中其它的一些例子出现得更早。

[注二] Fermat 猜想(现在被称为 Fermat 大定理) 的内容是： 方程 $x^n + y^n = z^n$ 在 n>2 时没有非零整数解。

[注三] 这让我想起了前文中提到的那个 Hardy 的电报。如果 Hardy 真的在那次航程中遇难了，他的那个电报也许就会象 Fermat 在页边上写下的那段话一样，成为数学史上的一个谜。

五. Riemann 的论文——零点分布与素数分布

在上节中我们看到，素数的分布与 Riemann ζ 函数之间存在着深刻的关联。这一关联的核心就是 J(x) 的积分表达式。由于 Riemann ζ 函数具有极为复杂的性质，这一积分同样也是极为复杂的。为了对这一积分做进一步的研究，Riemann 引进了一个辅助函数 ξ(s)[注一]：

$$\xi(s) = \Gamma(\frac{s}{2}+1)\ (s-1)\ \pi^{-s/2}\zeta(s)$$

引进这样的一个辅助函数有什么好处呢？可以证明，由上式定义的 ξ(s) 是一个整函数(Entire Function)，即在复平面上所有 s≠∞ 的点上都解析的函数。这样的函数在性质上

要比 Riemann ζ 函数简单得多，处理起来也容易得多。事实上，在所有非平庸的复变函数中，整函数是解析区域最为宽广的 (解析区域比它更大——即包括 $s=\infty$——的函数只有一种，那就是常数函数)。这是引进 $\xi(s)$ 的好处之一。利用这一辅助函数，我们在第一节 中提到的 Riemann ζ 函数满足的代数关系式

$$\zeta(s) = 2\Gamma(1-s)(2\pi)^{s-1}\sin(\pi s/2)\zeta(1-s)$$

可以表述为一个关于 s 与 1-s 对称的简单形式：

$$\xi(s) = \xi(1-s)$$

这是引进 $\xi(s)$ 的好处之二。

从 $\xi(s)$ 的定义中不难看到，$\xi(s)$ 的零点必定是 $\zeta(s)$ 的零点[注二]。由于我们已经知道，$\zeta(s)$在 $Re(s)>1$ 没有零点 (证明见 Euler 乘积公式 一文)，因此 $\xi(s)$ 在 $Re(s)>1$ 也没有零点；又由于 $\xi(s)=\xi(1-s)$，因此 $\xi(s)$ 在 $Re(s)<0$ 也没有零点。这表明 $\xi(s)$ 的所有零点都位于 $0\le Re(s)\le 1$ 的区域内。另一方面，$\zeta(s)$ 的零点除了平凡零点 $s=-2n$ (n 为自然数) 由于恰好是 $\Gamma(\frac{s}{2}+1)$ 的极点，从而不是 $\xi(s)$ 的零点外，全部都是 $\xi(s)$ 的零点，因此 $\xi(s)$ 的零点与 Riemann ζ 函数的非平凡零点重合。换句话说，$\xi(s)$ 将 Riemann ζ 函数的非平凡零点从全体零点中分离了出来，这是引进 $\xi(s)$ 的好处之三。将这一结果与上面刚刚证明的 $\xi(s)$ 的所有零点都位于 $0\le Re(s)\le 1$ 的区域内的结果联系起来，我们就得到了一个有关 Riemann ζ 函数零点分布的重要结果，那就是：Riemann ζ 函数的所有非平凡零点都位于复平面上 $0\le Re(s)\le 1$ 的区域内。这一结果虽然离 Riemann 猜想要求的所有非平凡零点都位于复平面上 $Re(s)=1/2$ 的直线上还相距甚远，但起码也算是万里长征的第一步。

Riemann 接着用 $\xi(s)$ 的零点对 $\ln\xi(s)$ 进行了分解：

$$\ln\xi(s) = \ln\xi(0) + \Sigma_\rho \ln(1-s/\rho)$$

其中 ρ 为 $\xi(s)$ 的零点 (也就是 Riemann ζ 函数的非平凡零点——这些家伙终于出场了！)。分解式中的求和对所有的 ρ 进行，并且是以先将 ρ 与 1-ρ 配对的方式进行的 (这一点很重要，因为上述级数是条件收敛的——但是在将 ρ 与 1-ρ 配对之后则是绝对收敛的)。这一分解式也可以写成等价的连乘积关系式：

$$\xi(s) = \xi(0)\ \Pi_\rho(1-s/\rho)$$

这样的连乘积关系式对于有限多项式来说是显而易见的 (只要 $\xi(0)\ne 0$)，但对于无穷乘积来说却绝非一目了然，它有赖于 $\xi(s)$ 是一个整函数这一事实，其完整证明直到 1893 年才由 Hadamard 在对整函数的无穷乘积表达式作系统研究时给出。 Hadamard 的这个证明是 Riemann 之后 34 年间在这一领域的唯一一个重要进展[注三]。

很明显，上述级数分解式的收敛与否与 $\xi(s)$ 的零点分布有着密切的关系。为 Riemann 研究了 $\xi(s)$ 的零点分布，并由此而提出了三个重要的命题：

1. 在 $0<Im(\rho)<T$ 的区间内，$\xi(s)$ 的零点数目大约为 $(T/2\pi)\ln(T/2\pi) - (T/2\pi)$。
2. 在 $0<Im(\rho)<T$ 的区间内， $\xi(s)$ 的位于 $Re(\rho)=1/2$ 的直线上的零点数目也大约为 $(T/2\pi)\ln(T/2\pi) - (T/2\pi)$。
3. $\xi(s)$ 的所有零点都位于 $Re(\rho)=1/2$ 的直线上。

在这三个命题中，第一个命题是为了证明级数分解式的收敛性所需要的(不过 Riemann 建立在这一命题基础上的说明——如我们在 [注三] 中所说——因过于简略，不足以构成证明)。对于这个命题 Riemann 的证明是指出在 $0<\text{Im}(\rho)<T$ 的区间内 $\xi(s)$ 的零点数目可以由 $d\xi(s)/2\pi i\xi(s)$ 沿矩形区域 $\{0<\text{Re}(\rho)<1, 0<\text{Im}(\rho)<T\}$ 的边界作路径积分得到。在 Riemann 看来，这点小小的积分算不上什么，因此他直接写下了结果 (即命题一)。Riemann 并且给出了该结果的相对误差为 $1/T$。但是 Riemann 显然大大高估了他的读者的水平，因为直到 46 年后 (1905 年)，他所写下的这一结果才由 von Mangoldt 给出了证明 (这一结果因此而被称为 Riemann-von Mangoldt 公式)。

不过 Riemann 留给读者们的这点智力挫折与他的第二个命题相比却又小巫见大巫了。将 Riemann 的第二个命题与前一个命题相比较可以看到，这第二个命题表明 $\xi(s)$ 的几乎所有的零点都位于 $\text{Re}(\rho)=1/2$ 的直线上。这是一个令人吃惊的命题，因为它比迄今为止——也就是 Riemann 的论文发表 144 年以来——人们在 Riemann 猜想上取得所有结果都要强得多!Riemann 在叙述这一命题的时候用的是完全确定的语气，这似乎表明，当他写下这一命题的时候，他认为自己对此已经有了证明。可惜的是他完全没有提及证明的细节，因此他究竟是怎么证明的 他的证明究竟是正确还是错误的？我们就无从得知了。 除了 1959 年的论文外，Riemann 还曾经在一封信件中提到过这一命题，他说这一命题可以从对 ξ 函数的一种新的表达式中得到，不过他还没有将之简化到可以发表的程度。这就是后人从 Riemann 留下的片言只语中得到的有关这一命题的全部信息。

Riemann 的这三个命题就象是三座渐次升高的山峰，一座比一座巍峨，攀登起来一座比一座困难.他的第一个命题让数学界等待了 46 年；他的第二个命题已经让数学界等待了 144 年； 而他的这第三个命题读者想必都看出来了，正是大名鼎鼎的 Riemann 猜想！它要让大家等待多久呢？没有人知道。但是据说著名的德国数学家 David Hilbert 有一次曾被人问到如果他能在 500 年后重返人间，他最想问的问题是什么？Hilbert 回答说他最想问的就是：是否已经有人解决了 Riemann 猜想[注四]？

正所谓山雨欲来风满楼，一直游刃有余、惯常在谈笑间让定理灰飞烟灭的 Riemann 到了表述这第三个命题——也就是 Riemann 猜想——的时候也终于一改举重若轻的风格，用起了象“非常可能”这样的不确定语气。Riemann 并且写道：“我们当然希望对此能有一个严格的证明，但是在经过了一些快速而徒劳的尝试后，我已经把对这种证明的寻找放在了一边，因为它对于我所研究的直接目标不是必须的”。Riemann 把证明放在了一边，整个数学界的心弦却被提了起来，直到今天还提得紧紧的。Riemann 猜想的成立与否对于 Riemann 的“直接目标”——即证明 $\ln\xi(s)$ 的级数分解式的收敛性——的确不是必须的 (因为那只要上述第一个命题就够了)，但对于今天的数学界来说却是至关重要的。粗略的统计表明， 在当今的数学文献中已经有超过一千条数学命题或“定理”以 Riemann 猜想的成立作为前提。Riemann 猜想的命运与提出这些命题或“定理”的所有数学家们的“直接目标”息息相关。另一方面，Riemann 对于 Riemann 猜想的表述方式也从一个侧面表明 Riemann 对于自己写下的命题是属于猜测性的还是肯定的是予以区分的。因此他对于那些没有注明是猜测性的命题——包括迄今无人能够证明的上述第二个命题——应该是有所证明的 (尽管由于他省略了证明， 我们无从知道那些证明是否正确)。

现在让我们回到对 $J(x)$ 的计算上来。利用 $\xi(s)$ 的定义及其分解式， 可以将 $\ln\zeta(s)$ 表示为：

$$\ln\zeta(s) = \ln\xi(0)+\Pi_\rho(1-s/\rho)-\ln\Gamma(\frac{s}{2}+1)-\frac{s}{2}\ln\pi-\ln(s-1)$$

对 $\ln\zeta(s)$ 作这样的分解目的是为了计算 $J(x)$，但是将这一分解式代入 $J(x)$ 的积分表达式后所得的各单项的积分并不都收敛， 因此 Riemann 在代入前先对 $J(x)$ 作了一次分部

积分，由此得到（读者可自行证明）：

$$J(x)=-\frac{1}{2\pi i}\frac{1}{\ln x}\int_{a-i\infty}^{a+i\infty}\frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{\ln\zeta(z)}{z}\right]x^z\,\mathrm{d}z$$

将 lnζ(s) 的分解式代入上式， 各单项可以分别积出， 其结果如下表所列：

lnζ(s) 分解式中的项	对应的积分结果
$-\ln(s-1)$	$\mathrm{Li}(x)$
$\Sigma_\rho \ln(1-s/\rho)$	$-\Sigma_{\mathrm{Im}(\rho)>0}\,[\mathrm{Li}(x^\rho)+\mathrm{Li}(x^{1-\rho})]$
$-\ln\Gamma(\frac{s}{2}+1)$	$\int_x^\infty \frac{\mathrm{d}t}{t(t^2-1)\ln t}$
$\ln\xi(0)$	$\ln\xi(0)=-\ln 2$
$-\frac{s}{2}\ln\pi$	0

在上述这些结果中，对 Σρ ln(1-s/ρ) 的积分最为复杂，其结果$-\Sigma_{\mathrm{Im}(\rho)>0}\,[\mathrm{Li}(x^\rho)+\mathrm{Li}(x^{1-\rho})]$是对级数逐项积分的结果。 这一结果是条件收敛的，不仅要如 lnξ(s) 的级数表达式中一样以将 ρ 与 1-ρ 配对的方式进行， 而且还必须依 Im(ρ) 从小到大的顺序求和。Riemann 在给出这一结果时承认逐项积分的有效性有赖于对 ξ 函数的“更严格”的讨论，但他说这是容易证明的。这一“容易证明”的结果在 36 年后（1895 年）被 von Mangoldt 所证明。另外值得指出的一点是， 在 Riemann 对这一级数的各单项进行积分的时候隐含了一个要求，那就是 0<Re(ρ)<1[注五]， 这比我们在前面已经证明的 0≤Re(ρ)≤1 要强。这一加强看似细微（不过是将等号排除掉而已）， 其实却——如我们在后文中将会看到的——是数论中一个非同小可的结果。Riemann 在文章中不仅没有对这一结果加以证明，连暗示性的说明也没有，这是他论文的一个漏洞。这个漏洞在 von Mangoldt 的证明中也同样存在[注六]。不过这一漏洞只是论证方法上的漏洞，是可以弥补的，论证的结果本身并不依赖于 0<Re(ρ)<1 这样的条件。由上面这些结果 Riemann 得到了 J(x) 的显形式：

$$J(x)=\mathrm{Li}(x)-\sum_{\mathrm{Im}\rho>0}\left[\mathrm{Li}(x^\rho)+\mathrm{Li}(x^{1-\rho})\right]+\int_x^\infty\frac{\mathrm{d}t}{t(t^2-1)\ln t}-\ln 2$$

这个结果，连同上节给出的 π(x) 与 J(x) 的关系式：

$$\pi(x)=\Sigma_n\,[\mu(n)/n]\,J(n^{1/n})$$

便是 Riemann 给出的素数分布的完整表达式，也是他 1859 年论文的主要结果。

Riemann 的这个结果给出的是素数分布的精确表达式，它的第一项（由 J(x)及 π(x) 的第一项共同给出）正是素数定理（当时还没有证明）所期待的结果 Li(x)。

但是 Riemann 的结果虽然给出了素数分布的精确表达式， 却没能直接证明远比该结果粗糙的素数定理，这是为什么呢？这其中的奥秘就在于 Riemann ζ 函数的非平凡零点，在于 J(x) 的表达式中那些与零点有关的项$-\Sigma_{\mathrm{Im}(\rho)>0}\,[\mathrm{Li}(x^\rho)+\mathrm{Li}(x^{1-\rho})]$：在 J(x) 的表达式中，所有其它的项都十分简单，也比较光滑，因此素数分布的细致规律——那些细致的疏密涨落——主要就蕴涵在了这一与 Riemann ζ 函数的非平凡零点有关的级数中。如上所述，这个级数是条件收敛的，这就是说它的收敛有赖与参与求和的各项——即来自不同零点的贡献——之间的相互抵消。这些来自不同零点的贡献就象一首盘旋起伏的舞曲，引导着素

数的细致分布。而这首舞曲的奔放程度——也就是这些贡献相互抵消的方式和程度——决定了素数分布与素数定理所给出的渐近分布之间的接近程度。所有这一切都定量地取决于 Riemann ζ 函数非平凡零点的分布。Riemann 给出的素数分布的精确结果之所以没能立即使对素数定理的直接证明成为可能，原因正是因为当时人们对 Riemann ζ 函数非平凡零点的分布还知道得太少(事实上当时人们所知道的也正是我们在上面已经证明的 $0 \le Re(\rho) \le 1$)， 无法有效地估计来自零点的那些贡献的大小， 从而也就无法有效地估计素数定理对素数实际分布——即 Riemann 的结果——的偏差。

那么 Riemann ζ 函数非平凡零点的分布对来自零点的那些贡献究竟有什么样的影响呢？数学家们已经取得了一系列结果。素数定理的证明本身就是其中一个， 我们将在后文中提及。在素数定理的证明之后，1901 年，瑞典数学家 von Koch (1870-1924) 进一步证明了，假如 Riemann 猜想成立，那么由素数定理给出的素数分布的绝对误差为 $O(x^{1/2}\ln x)$ (这是一个比素数定理更强的结果) 另一方面， 英国数学家 John Littlewood (1885-1977) 曾经证明， 素数定理给出的素数分布的绝对误差起码有 $Li(x^{1/2})\ln\ln\ln x$。这两者之间已经非常接近 (其主要项都是 $x^{1/2}$)。因此 Riemann 猜想的成立意味着素数的分布相对有序， 而假如 Riemann 猜想不成立， 假如 Riemann ζ 函数的部分非平凡零点偏离了 critical line，那么在素数的分布中就会出现紊乱， 素数定理对素数实际分布的偏差就会变大。对 Riemann 猜想的研究使数学家们看到了貌似随机的素数分布背后奇异的规律和秩序， 这种规律和秩序就体现在 Riemann ζ 函数的非平凡零点的分布之中， 它让数学家们目驰神移。

注释

[注一] Riemann 对 ξ 函数的定义与我们所用的略有差异，他的 ξ 函数用我们的 ξ 函数可以表示为 $\xi(s) = \xi(\frac{1}{2}+is)$。

[注二] 这是由于 Γ 函数没有零点，而 s-1 的唯一零点 s=1 又不是 ξ(s) 的零点 (因为 $\xi(1)=\xi(0)=-\zeta(0)=\frac{1}{2}$)。因此 ξ(s) 的零点只能出现在 ζ(s) 的零点处。

[注三] Riemann 虽然没有详细讨论上述无穷乘积表达式的证明，但他在写下与之等价的 lnξ(s) 的级数分解式之前提了一句： ξ(s) 是一个关于 $(s-\frac{1}{2})^2$ 的收敛极快的级数。这似乎暗示 ξ(s) 作为 $(s-\frac{1}{2})^2$ 的级数的收敛方式与它的无穷乘积表达式之间存在着联系。Hadamard 的证明确立了这种联系。此外，Riemann 通过讨论 ξ(s) 的零点分布对 lnξ(s) 的级数分解式的收敛性作了说明。虽然所有这些都因过于粗略，不足以构成证明，但这一暗一明两条思路后来都被证明是可以实现的。

[注四] 有意思的是 Hilbert 一度曾对 Riemann 猜想的解决抱有十分乐观的看法。他在 1919 年的一次演讲中表示在他自己的有生之年可望见到 Riemann 猜想的解决；在年轻听众的有生之年可望见到 Fermat 大定理的解决；而另一个问题——Hilbert 第七问题——才是最为困难的， 因为谁也没有希望看到它的解决。不料仅仅过了 10 年，Hilbert 就活着见到了他的第七问题的解决；75 年之后， Fermat 大定理也被解决了；而 Riemann 猜想却是谁也没能活着见到它的解决。

[注五] 确切地说是 $Re(\rho)>0$，但由于 ρ 与 1-ρ 总是同为零点，因此 $Re(\rho)>0$ 也意味

着 Re(ρ)<1。

[注六] 这里要区分两个不同的问题 一个是证明对级数可以进行逐项积分，另一个是计算级数各单项的积分。这个漏洞是出现后者之中的。

六. 错钓的大鱼

在 Riemann 的论文发表后的最初二三十年里，他所开辟的这一领域显得十分冷清，没有出现任何重大进展。如果把 Riemann 论文的全部内涵比做一座山峰的话，那么在最初这二三十年里数学家们还只在从山脚往半山腰攀登的路上，只顾星夜兼程、埋头赶路。那高耸入云的山颠还笼罩在一片浓浓的雾霭之中，真是高处不胜寒哪。但到了 1885 年，在这场沉闷的登山之旅中却爆出了一段惊人的插曲：有人忽然声称自己已经登顶归来！

这个人叫做 Thomas Stieltjes (1856-1894)，是一位荷兰数学家。1885 年，这位当时年方 29 岁的年青数学家在巴黎科学院发表了一份简报，声称自己证明了以下结果：

$$M(N) \equiv \Sigma_{n<N}\mu(n) = O(N^{1/2})$$

这里的 μ(n)是我们在 第四节 中提到过的 Möbius 函数，由它的求和所给出的函数 M(N) 被称为 Mertens 函数。这个命题看上去倒是面善得很：μ(n) 不过是一个整数函数，定义虽有些琐碎，却也并不复杂，M(N) 不过是对 μ(n) 的求和，证明它按照 $O(N^{1/2})$ 增长似乎不象是一件太困难的事情。但事实上这个其貌不扬的命题却是一个比 Riemann 猜想更强的结果！换句话说，证明了上述命题就等于证明了 Riemann 猜想(但反过来则不然， 否证了上述命题并不等于否证了 Riemann 猜想)。因此 Stieltjes 的简报等于是声称自己证明了 Riemann 猜想。

虽然当时 Riemann 猜想还远不象今天这么热门，消息传得也远不象今天这么飞快，但有人证明了 Riemann 猜想仍是一个非同小可的消息。别的不说，证明了 Riemann 猜想就等于证明了素数定理，而后者自 Gauss 等人提出以来折磨数学家们已近一个世纪之久，却还没能得到证明。与在巴黎科学院发表简报几乎同时，Stieltjes 给当时法国数学界的一位重量级人物 Charles Hermite (1822-1901)发去了一封信，重复了这一声明。但是无论在简报还是信件中 Stieltjes 都没有给出证明，他说自己的证明太复杂， 需要简化。

换作是在今天，一位年青数学家开的这样一张空头支票是很难引起数学界的反响的。但是十九世纪的情况却有所不同。因为当时学术界常有科学家做出成果却不公布(或只公布一个结果) 的事，Gauss 和 Riemann 都是此道中人。因此象 Stieltjes 那样声称自己证明了 Riemann 猜想，却不给出具体证明在当时并不算离奇。学术界的反应多少有点象现代法庭所奉行的无罪推定原则，即在出现相反证据之前倾向于相信声明成立。

但是相信归相信，数学当然是离不开证明的。因此大家就期待着 Stieltjes 发表具体的证明，其中期待得最诚心实意的当属 Hermite。Hermite 自 1882 就与 Stieltjes 保持着通信关系，直至 Stieltjes 十二年后过早去世为止。在这期间两人共交换了 432 封信件。Hermite 是当时复变函数论的大家之一，他与 Stieltjes 的关系堪称数学史上一个比较奇特的现象。Stieltjes 刚与 Hermite 通信时只是 Leiden 天文台的一名助理，而且就连这个助理职位还是靠了他父亲(Stieltjes 的父亲是荷兰著名的工程师兼国会成员)的关照。在此之前他在大学曾三度考试失败。好不容易进了天文台，Stieltjes 却〝身在曹营心在汉〞，干着天文观测的活，心里惦记的却是数学，并给 Hermite 写了信。照说当时一无学位、二无名声的 Stieltjes 要引起象 Hermite 这样的数学元老重视并不荣易，但 Hermite 是一位虔诚的天主教徒，他对数学怀有一种奇特的信仰，他相信数学存在是一种超自然的东西，寻常数学家只是偶尔才有机会了解数学的奥秘。那么什么样的人能够比〝寻常数学家〞更有机会了解数学的奥秘呢？Hermite 凭着自己的神秘主义眼光找到了一位，那就是默默无闻的观星之人 Stieltjes。Hermite 认为 Stieltjes 具有上帝赐于的窥视数学奥秘的眼光，他对

之充满了信任。在他与 Stieltjes 的通信中甚至出现了"你总是对的，我总是错的"这样极端的赞许。在这种神秘信仰与十九世纪数学氛围的共同影响下 Hermite 对 Stieljes 关于 Riemann 猜想的声明深信不疑。

但是无论 Hermite 如何催促，Stieljes 始终没有公布他的完整证明。一转眼五年过去了，Hermite 对 Stieljes 依然 "痴心不改"，他决定给对方来点 "利诱"。在 Hermite 提议下，法国科学院将 1890 年数学大奖的主题设为"确定小于给定数值的素数个数"。在 Hermite 看来，这个大奖将毫无悬念地落到他的朋友 Stieljes 的腰包里，因为这个大奖主题实质上就是证明素数定理，这比 Riemann 猜想弱得多。可惜直至大奖截止日期终了，Stieljes 依然毫无动静。

但是 Hermite 也没有完全失望，因为他的学生 Hadamard 提交了一篇论文，领走了大奖——肥水总算没有流入外人田。Hadamard 论文的主要内容正是我们在上节中提到的对 Riemann 论文中连乘积公式的证明。这一论文虽然离证明素数定理还有一定距离，却已足可获得大奖。几年之后，Hadamard 再接再励，终于一举证明了素数定理。Hermite 放出去的这根长线虽没能如愿钓到 Stieljes 及 Riemann 猜想，却错钓上了 Hadamard 及素数定理。斩获亦是颇为丰厚(当时素数定理其实比 Riemann 猜想更令数学界期待)。

那么 Stieljes 呢？没听过这个名字的读者可能会觉得他是一个浮夸无为的家伙，事实却不然。Stieljes 在分析与数论的许多方面都做出过重要的贡献。他在连分数方面的研究为他赢得了"连分数分析之父"的美誉，以他名字命名的 Stieljes 积分更是声名远播。但他的那份 Hardy 电报式的有关 Riemann 猜想的声明却终究没能为他赢得永久的悬念。现在数学家们普遍认为 Stieljes 关于 $M(N)=O(N^{1/2})$ 的证明是错误的，不仅如此，甚至连命题 $M(N)=O(N^{1/2})$ 本身是否成立也已经受到越来越多的怀疑[注一]。

七. 从零点分布到素数定理

素数定理自 Gauss 与 Legendre 以经验公式的形式提出(详见第三节) 以来，许多数学家对此做过研究。其中比较重要的结果是由俄国数学家 Pafnuty Chebyshev (1821-1894) 做出的。1850 年，Chebyshev 证明了对于足够大的 x，素数分布 $\pi(x)$ 与素数定理给出的分布 $Li(x)$ 之间的相对误差不超过 11%[注二]。

但在 Riemann 1859 年的工作以前，数学家们对素数定理的研究主要局限在实数域中。从这个意义上讲，即使撇开具体的结果不谈，Riemann 建立在复变函数基础上的工作仅就其方法而言也是对素数研究的一个重大突破。这一方法上的突破为素数定理的最终证明铺平了道路。

在第五节末尾我们曾提到，Riemann 对素数分布的研究之所以没能直接成为素数定理的证明，是因为人们对 Riemann ζ 函数非平凡零点的分布还知道得太少。那么为了证明素数定理，我们起码要知道多少有关非平凡零点分布的信息呢？这一点到了 1895 年随着 von Mangoldt 对 Riemann 论文的深入研究而变得明朗起来。von Mangoldt 的工作我们在第五节中已经提到过，正是他最终证明了 Riemann 关于 $J(x)$ 的公式。但 von Mangoldt 工作的价值比仅仅证明 Riemann 关于 $J(x)$ 的公式要深远得多。在他的研究中使用了一个比 Riemann 的 $J(x)$ 更简单有效的辅助函数 $\Psi(x)$，它的定义为：

$$\Psi(x) = \Sigma_{n<x} \Lambda(n)$$

其中 $\Lambda(n)$被称为 von Mangoldt 函数，它对于 $n=p^k$ (p 为素数，k 为自然数) 取值为 $\ln p$；对于其它 n 取值为 0。 运用 $\Psi(x)$，von Mangoldt 证明了一个本质上与 Riemann 关于 $J(x)$ 的公式等价的公式：

$$\Psi(x) = x-\Sigma_\rho(x^\rho/\rho)-\frac{1}{2}\ln(1-x^{-2})-\ln(2\pi)$$

其中有关 ρ 的求和与 Riemann 的 J(x) 中有关 ρ 的求和一样， 也是先将 ρ 与 1-ρ 配对，再依 Im(ρ)从小到大的顺序进行。

很明显，von Mangoldt 的 Ψ(x) 表达式比 Riemann 的 J(x) 简单多了。时至今日，Ψ(x) 在解析数论研究中差不多已完全取代了 Riemann 的 J(x)。引进 Ψ(x)的另一个重大好处是早在几年前，上文提到的 Chebyshev 就已经证明了：素数定理 π(x) ~ Li(x) 等价于 Ψ(x)~x(为了纪念 Chebyshev 的贡献，von Mangoldt 函数也被称为第二 Chebyshev 函数)。

将这一点与 von Mangoldt 的 Ψ(x)表达式联系在一起，不难看到素数定理成立的条件是 $\lim_{x\to\infty}\Sigma_\rho(x^{\rho-1}/\rho)=0$。但是要让 $x^{\rho-1}$ 趋于零，Re(ρ) 必须小于 1，换句话说 Riemann ζ 函数在直线 Re(s)=1 上必须没有非平凡零点。这就是我们为证明素数定理而必须知道的有关 Riemannζ 函数非平凡零点分布的信息[注三]。由于 Riemann ζ 函数的非平凡零点是以 ρ 与 1-ρ 成对的方式出现，因此这一信息也等价于 0<Re(ρ)<1。

读者们大概还记得，在第五节中我们曾经证明过 Riemannζ 函数的所有非平凡零点都位于 0≤Re(s)≤1 的区域内。

因此为了证明素数定理，我们所需知道的有关非平凡零点分布的信息要比我们已知的(也是当时数学家们已知的) 略多一些(但仍大大少于 Riemann 猜想所要求的)。这样，在经过了 Chebyshev、Riemann、Hadamard 和 von Mangoldt 等人的卓越努力之后，我们离素数定理的证明终于只剩下了最后一小步：即把已知的零点分布规律中那个可恶的等号去掉[注四]。这一小步虽也绝非轻而易举，却已难不住在 Riemann 峰上攀登了三十几个年头，为素数定理完整证明的到来等待了一个世纪的数学家们。von Mangoldt 的结果发表的第二年 (1896 年)，Hadamard 与比利时数学家 Charles de la Vallée-Poussin 就几乎同时独立地给出了证明，从而完成了 Gauss 以来数学界的一个重大心愿。那时 Stieljes 已经去世两年了。

经过素数定理的证明，人们对 Riemannζ 函数非平凡零点分布的了解又推进了一步，那就是：Riemann ζ 函数的所有非平凡零点都位于复平面上 0<Re(s)<1 的区域内。在 Riemann 猜想的研究中数学家们把这个区域称为 critical strip。

素数定理的证明——尤其是以一种与 Riemann 的论文如此密切相关的方式所实现的证明——让数学界把更多的注意力放到了 Riemann 猜想上来。四年后 (1900 年)的一个夏日，两百多位最杰出的数学家会聚到了巴黎，一位 38 岁的德国数学家走上了讲台，做了一次永载数学史册的伟大演讲。演讲的题目叫做“数学问题”，演讲者的名字叫做 David Hilbert，他恰好来自 Gauss 与 Riemann 的学术故乡——群星璀灿的 Göttingen 大学。他是 Göttingen 数学精神的伟大继承者，一位与 Gauss 及 Riemann 齐名的数学巨匠。Hilbert 在演讲中列出了二十三个对后世产生深远影响的数学问题，Riemann 猜想被列为其中第八个问题的一部分，从此成为整个数学界瞩目的难题之一。

二十世纪的数学大幕在 Hilbert 的演讲声中徐徐拉开，Riemann 猜想也迎来了一段新的百年征程。

注释

[注一]这是因为比 $M(N)=O(N^{1/2})$ 稍强、被称为 Mertens 猜想的命题：$M(N)<N^{1/2}$ 已于 1985 年被 Andrew Odlyzko 与 Herman te Riele 所否证。受此影响，目前数学家们倾向于认为 $M(N)=O(N^{1/2})$ 也并不成立，不过到目前为止还没人能够证明 (或否证) 这一点。

[注二] 比这更早一些，Chebyshev 还证明了：如果 $\lim_{x\to\infty} \{\pi(x)/[x/\ln(x)]\}$存在，它必定等于 1。Chebyshev 的研究对于 Riemann 的工作及后来人们对素数定理的证明都有影响。

[注三] 不过由于所处理的是无穷级数，严格的证明并不如我们叙述的那样简单。

[注四] 这也正是我们在第五节中提到的 Riemann 在计算 J(x) 过程中对与零点有关的级数进行单项积分时隐含的条件。

八. 零点在哪里？

随着 Riemann 论文中的外围命题——那些被 Riemann 随手写下却没有予以证明的命题——逐一得到证明，随着素数定理的攻克，也随着 Hilbert 演讲的聚焦作用的显现，数学界终于把注意力渐渐投向了 Riemann 猜想本身，投向了那座巍峨的主峰。

不知读者们有没有注意到，我们谈了这么久的 Riemann ζ函数，谈了那么久的ζ函数的非平凡零点，却始终没有谈及过任何一个具体的非平凡零点。这也是 Riemann 论文本身一个令人瞩目的特点：即它除了没有给所涉及的许多命题提供证明外，也没有给所提出的猜想提供数值计算方面的支持。Riemann 叙述了许多有关 ζ 函数非平凡零点的命题(比如第五节中提到的三大命题)，却没有给出任何一个非平凡零点的数值！

倘若那些非平凡零点是容易计算的，倒也罢了，可是就象被 Riemann 省略掉的那些命题个个都令人头疼一样，Riemann ζ函数的那些非平凡零点也个个都不是省油的灯。

它们究竟在哪里呢？

直到 1903 年(即 Riemann 的论文发表后的第 44 个年头)，丹麦数学家 Gørgen Gram 才首次公布了对 Riemann ζ函数前 15 个零点的计算结果[注一]。在这 15 个零点中，Gram 对前 10 个零点计算到了小数点后第六位，而后 5 个零点——由于计算繁复程度的增加——只计算到了小数点后第一位。为了让读者对 Riemann ζ函数的非平凡零点有一个具体的印象， 我们把这 15 个零点列在下面。与此同时，我们也列出了这 15 个零点的现代计算值(保留到小数点后第七位)，以便大家了解 Gram 计算的精度：

零点序号	Gram 的零点数值	现代数值
1	1/2 + 14.134725 i	1/2 + 14.1347251 i
2	1/2 + 21.022040 i	1/2 + 21.0220396 i
3	1/2 + 25.010856 i	1/2 + 25.0108575 i
4	1/2 + 30.424878 i	1/2 + 30.4248761 i
5	1/2 + 32.935057 i	1/2 + 32.9350615 i
6	1/2 + 37.586176 i	1/2 + 37.5861781 i
7	1/2 + 40.918720 i	1/2 + 40.9187190 i
8	1/2 + 43.327073 i	1/2 + 43.3270732 i
9	1/2 + 48.005150 i	1/2 + 48.0051508 i
10	1/2 + 49.773832 i	1/2 + 49.7738324 i
11	1/2 + 52.8 i	1/2 + 52.9703214 i

12	1/2 + 56.4 i	1/2 + 56.4462476 i
13	1/2 + 59.4 i	1/2 + 59.3470440 i
14	1/2 + 61.0 i	1/2 + 60.8317785 i
15	1/2 + 65.0 i	1/2 + 65.1125440 i

几十年来， 是数学家们第一次拨开迷雾实实在在地看到 Riemann ζ 数的非平凡零点，到那些蕴涵着素数分布规律的神秘家伙。 们都乖乖地躺在四十四年前 Riemann 出的那条奇异的critical line 上。Gram 的计算使用的是十八世纪三十年代发展起来的 Euler-Maclaurin 公式[注二]。 在只有纸和笔的年代里，这种计算是极其困难的，Gram 用了好几年的时间才完成对这 15 个零点的计算。但即便付出如此多的时间，付出极大的艰辛， 他在后五个零点的计算精度上仍不得不有所放弃。

在 Gram 之后，R. J. Backlund 于 1914 年把对零点的计算推进到了前 79 个零点。再往后，经过 Hardy、Littlewood 及 Hutchinson 等人的努力(包括计算方法上的一些改进)，到了 1925 年，人们已经知道了前 138 个零点的位置，它们都位于 Riemann 猜想所预言的 critical line 上。但是到了这个时候，建立在 Euler-Maclaurin 公式之上的计算已经复杂到了几乎难以逾越的程度。

九. Riemann 的手稿

随着数学界对Riemann 猜想兴趣的日益增加，这个猜想的难度也日益显露了出来。当越来越多的数学家在高不可测的 Riemann 猜想面前遭受挫折的时候，其中的一些开始流露出对 Riemann 1859 年论文的一些不满之意。我们在上面提到，Riemann 的论文既没有对它所提到的许多命题给予证明，又没有给出哪怕一个 ζ 函数非平凡零点的数值。尽管 Riemann 在数学界享有崇高的声誉，尽管此前几十年里人们通过对他论文的研究一再证实了他的卓越见解。但在攀登主峰的尝试屡遭受挫折，计算零点的努力又举步维艰的情况下，对 Riemann 的怀疑终于还是无可避免地出现了。

于是在承认 Riemann 的论文为“最杰出及富有成果的论文”之后 Laudau 开始表示：“Riemann 的公式远不是数论中最重要的东西，他不过是创造了一些在改进之后有可能证明许多其它结果的工具”；于是在为证明 Riemann 猜想度过一段“苦日子”之后 Littlewood 开始表示：“假如我们能够坚定地相信这个猜想是错误的，日子会过得更舒适些”；于是就连胆敢用 Riemann 猜想跟上帝耍计谋的 Hardy 也开始认为 Riemann 有关零点的猜测只不过是个猜测而已，nothing more。“Nothing more”的意思便是纯属猜测，没有任何计算及证明依据。换句话说数学家们开始认为 Riemann 论文中的一切大致也就是他在这一论题上所做过的一切，他的猜想其依据的只是直觉，而非证据。

那么 Riemann 猜想究竟只是凭借直觉呢还是有着其它的依据？Riemann 的论文究竟是不是他在这方面的全部研究呢？既然 Riemann 的论文本身没有为这些问题提供线索，答案自然就只能到他的手稿中去寻找了。

我们曾经提到， 在 Riemann 那个时代许多数学家公开发表的东西往往只是他们所做研究的很小一部分，因此他们的手稿及信件就成为了科学界极为珍贵的财富。这种珍贵绝不是因为如今人们习以为常的那种名人用品的庸俗商业价值，而是在于其巨大的学术价值。因为通过它们，人们不仅可以透视那些伟大先辈们的“Beautiful Mind”，更可以挖掘他们未曾发表过的研究成果，那是一种无上的宝藏。

不幸的是，Riemann 手稿的很大一部分却在他去世之后被他可恶的管家付之了一炬，只有一小部分被他妻子 Elise 抢救了出来。Elise 把那些劫后余生的数学手稿大部分交给了

Riemann 生前的挚友、数学家 Richard Dedekind (1831-1916)。但是几年之后，Elise 又后悔了，因为她觉得那些数学手稿中还夹带着一些私人及家庭的信息，于是她向 Dedekind 索回了一部分手稿。在这部分手稿中，有许多几乎通篇都是数学，只在其中夹带了极少量的私人信息，比如一位朋友的姓名等，其中更有一本小册子是 Riemann 1860 年春天在巴黎时的记录。那正是他发表有关 Riemann 猜想的论文后的几个月。那几个月巴黎的天气十分糟糕，很多时候 Riemann 都待在住所里研究数学。许多人猜测，在那段时间里 Riemann 所思考的很可能与他几个月前研究的 Riemann ζ 函数有关联，因此那本被 Elise 索回的小册子中很可能记录了与 Riemann 猜想有关的一些想法。可惜那本数学家们非常渴望获得的小册子从此就再也没有出现过，直到今天，它的去向依然是一个谜。有人说它曾被德国数学及数学史学家 Erich Bessel-Hagen(1898-1946)获得过，但是 Bessel-Hagen 死于二战后的混乱年月中，他的遗物始终没有被人找到过。

那些有幸躲过管家的火把、又没有被 Elise 索回的手稿，Dedekind 将它们留在了 Göttingen 大学图书馆，那就是数学家和数学史学家们可以看到的 Riemann 的全部手稿 (Nachlass)。

自 Riemann 的手稿存放在 Göttingen 大学图书馆以来，陆续有一些数学家及数学史学家前去研究。但是只要想一想 Riemann 正式发表的有关 Riemann 猜想的论文尚且如此艰深，就不难想象研读他那些天马行空、诸般论题混杂、满篇公式却几乎没有半点文字说明的手稿该是一件多么困难的事情。许多人满怀希望而来，却又两手空空、黯然失望而去。

Riemann 的手稿就象一本高明的密码本，牢牢守护着这位伟大数学家的思维奥秘。

但是到了 1932 年，终于有一位数学家从那些天书般的手稿中获得了重大的发现！这一发现一举粉碎了那些认为 Riemann 的论文只有直觉而无证据的猜测，并对 Riemann ζ 函数非平凡零点的计算方法产生了脱胎换骨般的影响，让在第 138 个零点附近停滞多年的 Euler-Maclaurin 方法相形见拙。这一发现也将它的发现者的名字与伟大的 Riemann 联系在了一起，从此不朽。

这位破解天书的发现者叫做 Carl Siegel(1896-1981)，他是 Riemann 的同胞——一位德国数学家。

十. 探求天书

Siegel 是一位非常反战的德国人，早年曾因拒服兵役而遭拘压，幸亏 Landau 的父亲出面帮助才得重归自由。他曾计划在柏林学习天文学，因为天文学是看上去最远离战争的学科。 但是入学那年的天文学课程开得较晚，为了打发时光，他去听了 Georg Frobenius (1849-1917) 数学课，这一听很快改变了他的人生旅途，他最终成为了一名数学家。

Siegel 于 1919 年来到 Göttingen,跟随 Landau 研究数论。当时 Hilbert 的二十三个数学问题已经非常出名，而 Landau 本人对 Riemann 猜想也颇有研究，在这种环境的影响下，Siegel 也开始了对 Riemann 猜想——Hilbert 第八问题的一部分——的研究。他对 Riemann 猜想的一些想法得到了 Hilbert 本人的赏识，在 Hilbert 的支持下，Siegel 于 1922 年获得了 Frankfurt 大学的教职。

但尽管如此，Siegel 对 Riemann 猜想的研究并没有取得突破性的进展。正当他为此苦恼的时候，一封来自数学及数学史学家 Bessel-Hagen 的信寄到了他的案头。Bessel-Hagen 当时正在研究 Riemann 的手稿，但和 Siegel 研究 Riemann 猜想一样苦苦得不到进展。由于 Bessel-Hagen 自身的背景侧重于数学史，对于破解 Riemann 的手稿来说这样的背景显然还嫌不够，于是他想邀请纯数学家来试试，看看他们是否能有所突破。Göttingen 的数学家中对 Riemann 猜想感兴趣的当首推 Hilbert 和 Landau，但这两位都是大师级的人物，Bessel-Hagen 自不敢贸然相扰，于是他把目光投向了正在研究

Riemann 猜想的 Siegel，邀请他来研究 Riemann 的手稿。

对 Siegel 来说 Bessel-Hagen 的邀请不失为一个散心的机会。另一方面，如我们在上节所说，当时数学界对 Riemann 及其猜想的怀疑已经开始蔓延，这种氛围也影响到了 Göttingen，Riemann 是不是真的只凭直觉提出他的猜想？这也是 Siegel 有意一探究竟的谜团。于是 Siegel 写信向 Göttingen 图书馆索来了 Riemann 的手稿。

当那位已被岁月涂抹成只凭直觉研究数学的前辈宗师的手稿终于出现在 Siegel 眼前的时候，他不由地想起了 Gauss 爱说的一句话：工匠总是会在建筑完成后把脚手架拆除的。现在他所看到的正是一位最伟大工匠的脚手架，任何人只要看上一眼就绝不会再相信那些有关 Riemann 只凭直觉研究数学的传言。只可惜那些散布传言的数学家们——包括与 Riemann 手稿近在咫尺的睿智的 Göttingen 数学家们——竟然谁也没有费心来看一眼这些凝聚着无比智慧的手稿！

在 Riemann 的手稿中，Siegel 发现了 Riemann 论文中只字未提的 Riemann ζ 函数的前三个零点的数值[注一]！很显然，这表明 Riemann 的论文背后是有着计算背景的。Riemann 的这一计算比我们在第八节中提到的 Gram 的计算早了 44 年。这倒也罢了，因为 Gram 对零点的计算虽比 Riemann 的晚，但精度却比 Riemann 的高得多。但是 Siegel 对 Riemann 计算零点的方法进行了细致的整理研究，却吃惊地发现 Riemann 所用的方法不仅远远胜过了 Gram 所用的 Euler-Maclaurin 公式，也远远胜过了 Hardy 和 Littlewood 对 Euler-Maclaurin 公式的改进。一句话，Riemann 用来计算零点的方法远远胜过了数学界已知的任何方法！当时已 1932 年，距离 Riemann 猜想的提出已有 73 个年头，距离 Riemann 逝世也已有 66 个年头，Riemann 又一次跨越时间远远地走到了整个数学界的前面。而且 Riemann 的这一公式是如此的复杂[注二]，有些数学家甚至认为假如不是 Siegel 把它从 Riemann 的手稿中整理出来的话，也许直到今天，数学家们都无法独立地发现它。

Siegel 在整理这一公式上的功绩和所付出的辛劳是怎么评价也不过分的，如我们在 上节 中所说，Riemann 的手稿上诸般论题混杂、满篇公式却几乎没有半点文字说明。而且 Riemann 晚年的生活很不宽裕，用纸十分节约，每张稿纸的角角落落都写满了东西，使得整个手稿更显混乱。再加上 Riemann 所写的那些东西本身的艰深。Siegel 能从中整理出如此复杂的公式对数学界实是功不可没，为了表达对 Siegel 工作的敬意，数学家们把这一公式称为 Riemann-Siegel 公式。Siegel 一生对数学多有贡献，但其中最杰出的一项也许就是这一公式。Riemann 若泉下有知，也当乐见他的这位后辈同胞的名字通过这一公式与自己联系在一起，因为在这之后，再也没有人会怀疑他论文背后的运算背景了。

发表于 1932 年的 Riemann-Siegel 公式是 Göttingen 数学辉煌的一抹余辉。随着纳粹在德国日益横行，曾经是数学圣地的 Göttingen 一步步地走向了衰落。1933 年，Landau 因其“犹太式的微积分与雅里安 (Aryan) 的思维方式背道而驰”被剥夺了授课资格，离开了他一生挚爱的数学讲堂。出于对战争的厌恶，Siegel 于 1940 年离开了德国。Göttingen 的衰落是德国文化史上最深重的悲剧之一。在这场悲剧中最痛苦的也许要算是 Hilbert，他是自 Gauss 和 Riemann 之后 Göttingen 数学传统的灵魂人物，从某种意义上讲，Göttingen 也是 Hilbert 的灵魂。他一生为发扬 Göttingen 的数学传统尽了无数的心力，Göttingen 记录了他一生的荣耀和自豪，而今在他年逾古稀的时候却要残酷地亲眼目睹这一切的辉煌烟消云散。1943 年，Hilbert 黯然离开了人世，Göttingen 的一个时代走到了终点。

十一. Riemann-Siegel 公式

Riemann-Siegel 公式的推导极其复杂，不可能在本文中加以介绍。不过我们将简单

叙述一下计算 Riemann ζ函数非平凡零点的基本思路，并给出 Riemann-Siegel 公式的表达式，以便读者有一个大致的了解。

读者也许还记得，在第五节中我们曾引进过一个辅助函数

$$\xi(s) = \Gamma(\frac{s}{2} + 1)\ (s-1)\ \pi^{-s/2}\,\zeta(s)$$

它的零点与 Riemann ζ函数的非平凡零点重合。因此，我们可以通过对ξ(s)零点的计算来确定 Riemann ζ函数的非平凡零点。这是计算 Riemann ζ函数零点的基本思路。由于ξ(s)满足一个特殊的条件：ξ(s)=ξ(1-s)，运用复变函数论中的反射原理(reflection principle)很容易证明(读者不妨自己试试)，在 Re(s)=1/2 的直线(即 Riemann 猜想中的 critical line)上ξ(s)的取值为实数。因此在 critical line 上通过研究ξ(s)的符号改变就可以确定零点的存在。这是利用ξ(s)计算零点的一个极大的优势。在下文中我们将只考虑 critical line 上的情形，为此令 s=1/2+it。利用 ξ(s) 的定义可以证明 (请读者自行完成)：

$$x(\frac{1}{2}+it) = [e^{\mathrm{Re}\ln\Pi(\frac{s}{2}-1)}p^{-\frac{1}{4}}(-t^2-\frac{1}{4})/2][e^{i\,\mathrm{Im}\Pi(\frac{s}{2}-1)}p^{-\frac{it}{2}}z(\frac{1}{2}+it)]$$

很明显，上式中第一个方括号内的表达式始终为负，因此在计算ξ(s) 的符号改变——从而确定零点——时可以忽略。因此要想确定 Riemann ζ函数的零点，只需研究上式中第二个方括号内的表达式就可以了。我们用 Z(t)来标记这一表达式，即：

$$Z(t) = e^{i\,\mathrm{Im}\Pi(\frac{s}{2}-1)}p^{-\frac{it}{2}}z(\frac{1}{2}+it)$$

至此，研究 Riemann ζ函数在 critical line 上的零点就归结为研究 Z(t)的零点，而后者又可以归结为研究 Z(t)的符号改变。

Riemann-Siegel 公式就是关于 Z(t)的渐进展开式，它可以表示为：

$$Z(t) = 2\sum_{n^2<\frac{t}{2p}} n^{-\frac{1}{2}}\cos[q(t) - t\log n] + R(t)$$

其中：

$$\theta(t) = \frac{t}{2}\log\frac{t}{2\pi} - \frac{t}{2} - \frac{\pi}{8} + \frac{1}{48t} + \frac{7}{5760t^3} + \ldots$$

$$R(t) \sim (-1)^{N-1}\left(\frac{t}{2\pi}\right)^{-1/4}\left[C_0 + C_1\left(\frac{t}{2\pi}\right)^{-1/2} + C_2\left(\frac{t}{2\pi}\right)^{-2/2} + C_3\left(\frac{t}{2\pi}\right)^{-3/2} + C_4\left(\frac{t}{2\pi}\right)^{-4/2}\right]$$

上面式子中的 R(t)被称为剩余项(reminder)，其中的 N 为$(t/2\pi)^{1/2}$的整数部分，R(t)中各项的系数分别为：

$$C_0 = \Psi(p) \equiv \frac{\cos\left[2\pi\left(p^2 - p - \frac{1}{16}\right)\right]}{\cos(2\pi p)}$$

$$C_1 = -\frac{1}{2^5\cdot 3\cdot \pi^2}\Psi^{(3)}(p)$$

$$C_2 = \frac{1}{2^{11} \cdot 3^2 \cdot \pi^4} \Psi^{(6)}(p) + \frac{1}{2^6 \cdot \pi^2} \Psi^{(2)}(p)$$

$$C_3 = -\frac{1}{2^{16} \cdot 3^4 \cdot \pi^6} \Psi^{(9)}(p) - \frac{1}{2^8 \cdot 3 \cdot 5 \cdot \pi^4} \Psi^{(5)}(p) - \frac{1}{2^6 \cdot \pi^2} \Psi^{(1)}(p)$$

$$C_4 = \frac{1}{2^{23} \cdot 3^5 \cdot \pi^8} \Psi^{(12)}(p) + \frac{11}{2^{17} \cdot 3^2 \cdot 5 \cdot \pi^6} \Psi^{(8)}(p) + \frac{19}{2^{13} \cdot 3 \cdot \pi^4} \Psi^{(4)}(p) + \frac{1}{2^7 \cdot \pi^2} \Psi(p)$$

其中 p 为$(t/2\pi)^{1/2}$的分数部分，$\Psi(n)^{(n)}$为 $\Psi(p)$ 的 n 阶导数。

这就是 Siegel 从 Riemann 手稿中整理出来的计算 Riemann ζ 函数零点的公式[注三]。确切地讲它只是计算 Riemann ζ 函数数值的公式，要想确定零点的位置还必须通过多次计算逐渐逼近，其工作量比单单计算 Riemann ζ 函数的数值大得多。读者也许会感到奇怪，如此复杂的公式加上如此迂回的步骤，在没有计算机的年代里能有多大用处？的确，计算 Riemann ζ 函数的零点即使使用 Riemann-Siegel 公式也是极其繁复的，别的不说，只要看看 C_4 中对$\Psi(p)$的导数竟高达 12 阶之多就足令人头疼了。但是同样一件工作，在一位只在饭后茶余瞥上几眼的过客眼里与一位对其倾注生命、不惜花费时光的数学家眼里，它的可行性是完全不同的。就象在一位普通人、甚或是一位普通数学家的眼里 Riemann 能做出如此深奥的数学贡献是不可思议的一样。

不过，也不要把 Riemann-Siegel 公式看得太过可怕，因为在下一节中，我们就将一起动手用这一公式计算一个 Riemann ζ 函数的非平凡零点。当然，我们会适当偷点懒，也会用用计算器，甚至还要用点计算机软件。毕竟，我们与 Siegel 之间又隔了七十多个年头，具备了偷懒所需的信息和工具。然后，我们将继续我们的旅途，去欣赏那些勤奋的人们所完成的工作，那才是真正的风景。

注释

[注一] 后来的一些数学史学家甚至认为 Riemann 可能计算过多达 20 个零点。

[注二] 当然这种复杂性指的是推导上的复杂，而不是用来计算零点时的复杂——后者虽然也很复杂，却比传统的 Euler-Maclaurin 公式来得简单。

[注三] 有两点需要提醒读者：一是 Riemann 手稿中 C_4 中$\Psi(p)$ 的系数与 Siegel 给出的不同；二是我们没有使用 Siegel 原始论文中的记号。

十二. 休闲课题：围捕零点

听说时下流行一种休闲方式叫做 DIY (Do It Yourself)，讲究自己动手做一些原本只有工匠才做的事，比方说自己动手做件陶器什么的。在象我这样懒散的人看来这简直比工作还累，可如今许多人偏偏就兴这个，或许是领悟了负负得正(累累得闲？)的道理吧。既是大势如此，我们也乐得共襄盛举，安排“休闲”一下，让大家亲自动手用 Riemann-Siegel 公式来计算一个 Riemann ζ 函数的非平凡零点。

DIY 一般有个特点，那就是课题虽然选得颇见难度，做起来通常却是挑最简单的来做，以免打击休闲的积极性。我们计算零点也一样，挑相对简单的零点来计算。那么什么样的零点比较容易计算呢？显然是那些听 Riemann 的话，乖乖地躺在 critical line 上的零点——因为否则的话 Riemann 猜想早被推翻了。

在 Riemann-Siegel 公式中有许多复杂的东西，其中最令人头疼的是求和，因为它使计算量成倍地增加。但幸运的是那个求和是对 $n^2<(t/2\pi)$ 进行的，因此如果 $t<8\pi\approx25$，求和就只有 $n=1$ 一项。这显然是比较简单的，因此我们狡猾的目光就盯在了这一区间上。在这一区间上，Riemann-Siegel 公式简化成为：

$$Z(t) = 2\cos[\theta(t)] + R(t)$$

在正式围捕之前，我们先做一点火力侦察——粗略地估计一下猎物的位置。我们要找的是使 $Z(t)$为零的点，直接寻找显然是极其困难的，但我们注意到 $2\cos[\theta(t)]$ (通常被称为主项) 在 $\theta(t)=(m+1/2)\pi$ 时为零(m 为整数)，这是一个不错的出发点。由上节中 $\theta(t)$ 的表达式不难证明，在所有这些使 $2\cos[\theta(t)]$ 为零的 $\theta(t)$ 中， $\theta=-\pi/2$ (即 $m=-1$) 是使 t 在 $0<t<25$ 中取值最小的，它所对应的 t 为 $t\approx14.5$。这是我们关于零点的第一个估计值。 纯以数值而论，它还算不错，相对误差约为百分之三。

接下来我们对这个估计值进行一次修正。修正的理由是显而易见的，因为 $t\approx14.5$ 时 $R(t)$ 明显不为零。为了计算 $R(t)$ 我们注意到 $t\approx14.5$ 时$(t/2\pi)^{1/2}\approx1.5$，因此 $R(t)$ 中的参数 N 为 1，p [$(t/2\pi)^{1/2}$的分数部分]约为 0.5。由此可以求出 $R(t)$中的第一项——$C_0(t/2\pi)^{-1/4}$——约为 0.3。

为了抵消这额外的 0.3，我们需要对 t 进行修正，使 $2\cos[\theta(t)]$ 减少 0.3。我们采用线性近似$\Delta t\approx\Delta F(t)/F'(t)$来计算这一修正值。为此注意到 $2\cos[\theta(t)]$在 $t\approx14.5$ 处的导数为 $-2\theta'(t)\sin[\theta(t)]\approx-2(1/2)\ln(14.5/2\pi)\sin(-\pi/2)\approx0.83$。由此可知 t 需要修正为 $t+\Delta t\approx14.5-0.3/0.83\approx14.14$。这个数值与零点的实际值之间的相对误差仅为万分之四。但是需要提醒读者的是，这种估计——无论它多高明——都不足以证明零点的存在，它至多只能提供一个围捕零点的范围。

那么究竟怎样才能证明零点的存在呢？我们在上节已经提供了方法。那就是通过计算 $Z(t)$ 的符号，如果 $Z(t)$ 在某两点的符号相反，就说明 Riemann ζ 函数在这两点之间内存在零点。我们上面所做的估计就是为这一计算做准备的。现在我们就来进行这样的计算。由于我们已经发现在 $t=14.14$ 附近可能存在零点，因此我们在 $14.1\leqslant t\leqslant14.2$ 的区间上撒下一张小网。如果我们的计算表明 $Z(t)$在这一区间的两端，即 $t=14.1$ 与 $t=14.2$ 具有不同的符号，那就证明了 Riemann ζ函数在 $t=14.1$ 与 $t=14.2$ 之间存在零点[注一]。下面我们就来进行计算：

对于 $t=14.1$，$(t/2\pi)^{1/2}\approx1.498027$， $\theta(t)\approx-1.742722$。因而主项 $2\cos[\theta(t)]\approx-0.342160$，剩余项 $R(t)$中 $p\approx0.498027$，从而其中第一项(C_0 项) $C_0(t/2\pi)^{-1/4}\approx0.312671$。由这两部分(即主项及剩余项中的第一项)可得：

$$Z(14.1)\approx-0.342160 + 0.312671=-0.029489$$

类似地，对于 $t=14.2$，$(t/2\pi)1/2\approx1.503330$， $\theta(t)\approx-1.702141$。因而主项 $2\cos[\theta(t)]\approx-0.261934$，剩余项 $R(t)$中 $p\approx0.503330$，从而其中第一项(C_0 项) $C_0(t/2\pi)^{-1/4}\approx0.312129$。由这两部分(即主项及剩余项中的第一项)可得：

$$Z(14.2) \approx -0.261934 + 0.312129 = 0.050195$$

显然，如我们所期望的，$Z(14.1)$ 与 $Z(14.2)$ 符号相反，这表明在 $t=14.1$ 与 $t=14.2$ 之间存在 Riemann ζ函数的零点。当然，我们还没有考虑 C_1~C_4 项。这些项中带有 C_0 的各阶导数，计算起来工作量非同小可，有违休闲的目的，因此就不费心了。熟悉计算软件的读者可以用 Mathematica、Maple 或 Matlab 一类的工具来算一下。我们把所有这些计算结果都列在下表中：

	t=14.1	t=14.2
N	1	1

p	0.498027	0.503330
θ(t)	-1.742722	-1.702141

2cos[θ(t)]	-0.342160	-0.261934
C_0 项	0.312671	0.312129
C_1 项	0.000058	0.000097
C_2 项	0.001889	0.001872
C_3 项	0.000001	0.000002
C_4 项	0.000075	0.000074
Z(t)	-0.027446	0.052042

从这些结果中可以看到，剩余项中的高阶项的贡献虽然有所起伏，但与第一项相比总体上很小。对于我们来说，这显然是很幸运的结果，因为否则的话，我们就得休闲不成反卖苦力了。这还是 t 较小的情况。随着 t 的增加，由于高阶项中所含 t 的负幂次较高，其贡献会变得越来越小[注二]，但要严格表述这种趋势并予以证明，却绝非轻而易举。事实上 Riemann-Siegel 公式作为 Z(t) 的渐进展开式，其敛散性质与误差估计都是相当复杂的。

现在我们知道了 Riemann ζ函数在 t=14.1 与 t=14.2 之间存在零点。如果我们再仔细点，注意到 Z(14.1) 与 Z(14.2)距离 Z(t)=0 的远近之比为 0.027446:0.052042，用线性内插法可以推测零点的位置为

$$t \approx 14.1+(14.2-14.1)\times 0.027446/(0.027446+0.052042) \approx 14.1345$$

这与现代数值 t=14.1347 的相对偏差只有不到十万分之二！即使只估计到 C_0 项（这是我们自己动手所及的范围），其误差也只有不到万分之二。

好了，猎物在手，我们的简短休闲也该见好就收了。大家是否觉得有点成就感呢？要知道，Riemann ζ函数的零点可是在 Riemann 的论文发表之后隔了四十四年才有人公布计算结果的哦。当然，我们用了 Riemann-Siegel 公式，但这没什么，一个好汉三个帮嘛，再说了， DIY 哪有真的百分之百从头做起，连工具设备都包括在内的？想象一下，如果你 DIY 出来的陶器能够把缺陷控制在万分之二以内，那是何等的风光？当然，倘若你可以退回一百多年，把这个结果抢在 Gram 之前公布一下，那就更风光了。

在本节最后，还有一件可能让大家有成就感的事要提一下。那就是我们所用的估计零点的方法——即从使 2cos[θ(t)] 为零的点出发，然后依据 R(t)的数值对其进行修正 [注三]，最后用 Z(t)的符号来确定零点的存在，暗示 Riemann ζ函数在 critical line 上的零点数目大致与 cos[θ(t)] 的零点数目相当。而后者大约有（请大家 DIY）$\theta(t)/\pi \sim (t/2\pi)\ln(t/2\pi)-(t/2\pi)$ 个。不知大家是否还记得，这正是我们在第五节中介绍过的 Riemann 的三个命题中迄今无人能够证明的第二个命题！当然，我们这个也不是证明（真可惜，否则的话，嘿嘿……），但这应该使大家对我们休闲手段之高明有所认识吧？

注释

[注一] 要注意的是，Z(t)在一个区间的两端具有不同符号只是 Riemann ζ函数在该区间存在零点的充分条件，而非必要条件。换句话说，假如我们不幸发现 Z(t) 在我们所取的两点上具有相同的符号，不能直接得出结论说 Riemann ζ 函数在这两点之间不存在零点。至于这是为什么，请大家 DIY。

[注二]但另一方面，随着 t 的增加，Riemann-Siegel 公式中的求和所包含的项数会逐渐增加 因此计算的总体复杂程度并不呈现下降趋势。

[注三]对于求和中有不止一项的情形，修正所依据的不仅仅是 R(t)，但思路是类似的。

十三. 从纸笔到机器

Riemann-Siegel公式的发表大大推进了人们对Riemann ζ函数非平凡零点的计算。如我们在前两节所看到的，Riemann-Siegel 公式中的求和项数是由 $n^2<(t/2\pi)$ 确定的，这表明用 Riemann-Siegel 公式计算一个位于 s=1/2+it 附近的零点所需的计算量为 $O(t^{1/2})$。而在这之前人们所用的 Euler-Maclaurin 公式计算同样的零点所需的计算量约为 O(t)。这两者的差别——也就是 Riemann-Siegel 公式相对于 Euler-Maclaurin 公式的优越之处——随着 t 的增大而变得越来越明显。

Riemann-Siegel 公式发表大约四年后，Hardy 的学生、英国数学家 Edward Titchmarsh(1899-1963)成功地计算出了 Riemann ζ函数前 1041 个零点的位置，它们全都位于 critical line 上。这是十一年来数学家们首次突破我们在第八节提到过的由 Hutchinson 创造的 138 个零点的记录。Titchmarsh 的工作在 Riemann ζ函数非平凡零点计算史上的地位是双重的：从计算方法上讲，它是数学家们首次用Riemann-Siegel 公式取代 Euler-Maclaurin 公式进行大规模零点计算；从计算手段上讲，Titchmarsh 的计算使用了英国海军部用来计算天体运动及潮汐的一台打孔式计算机(punched-card machine)，这是数学家们在零点计算上首次用机器计算取代传统的纸笔计算。这两个转折是数学与技术相辅相成的结果，它奠定了直到今天为止人们对 Riemann ζ函数非平凡零点进行计算的基本模式。

Titchmarsh 之后零点的计算因第二次世界大战的爆发中断了十几年。战后最先将计算推进下去的是著名的英国数学家 Alan Turing(1912-1954)。Turing 其实早在战前就对 Riemann 猜想产生了兴趣。与当时许多其他年轻数学家一样，Turing 对 Hilbert 的数学问题很感兴趣，这其中又尤以第十问题与第八问题(Riemann 猜想是第八问题的一部分)最让他着迷[注一]。他后来主要的研究都是以这两个问题为主轴展开的。1936 年 Turing 到 Princeton 大学读研究生，在那里见到了来访的 Hardy (他原本希望能在 Princeton 见到 Gödel，可惜后者当时已经去了欧洲)。那时 Hardy 对 Riemann 猜想的态度已经相当悲观。 这种悲观情绪对 Turing 产生了影响，他觉得这么多年来所有证明 Riemann 猜想的努力都归于失败，也许是到了换个角度思考问题的时候了。人们一直始终无法证明 Riemann 猜想，也许并非因为它太难，而是因为它根本就不成立！

一个数学命题，它的成立需要证明，不成立同样需要证明。假如 Riemann 猜想真的不成立，我们怎样才能证明这一点呢？我们当然可以试图从数学上直接证明其不成立，这是一种方法。但还有一种办法，那就是找到一个反例，即找到一个不在 critical line 上的零点。这种方法的好处是不在乎多少，只要一个反例就足够了。被后世誉为 “计算机与人工智能之父”的 Turing 显然对后一种方法情有独钟。当时 Turing 已经提出了后来以他名字命名的 Turing 机的概念。很自然的，他希望建造一台机器来计算零点。但是这一工作起步不久，英国就卷入了二战，Turing 开始参与英国情报部门破译德军密码的工作，建造机器的计划被搁置了下来。战争结束后，Turing 渐渐恢复了建造机器及计算零点的计划。Turing 虽然是以其对计算机及人工智能领域的卓越贡献著称的，但他在传统数学领域也有相当深厚的功力，早在读本科的时候，他就曾独立证明了概率论中著名的中心极限定理 (可惜比 J. W. Lindeberg 晚了十余年)。在建造机器的同时，Turing 对计算零点的数学方法也进行了研

究，并做了一些改进。

经过几年的努力，到了二十世纪五十年代初，Turing 终于完成了他的机器，并且比创造战前记录的 Titchmarsh 略进一步，计算出了前 1104 个零点。不过他试图寻找 Riemann 猜想反例的努力并不成功，因为所有这些零点全部位于 critical line 上，Riemann 猜想在他计算所及的范围内岿然不动。在那之后，Turing 的机器坏掉了。几乎与此同时，他的个人生活也遭遇了极大的挫折。他于 1952 年被控犯有当时属于违法的同性恋行为，受到强制药物治疗及缓刑的处罚。两年后他被发现因氰化物中毒死于寓所。多数人相信他是自杀。[注二]

在 Turing 之后，随着计算机发展的加速，数学家们对零点的计算也越来越快。1956 年，D. H. Lehmer 计算了前 25000 个零点；两年后 N. A. Meller 把这一记录推进到了 35337 个零点；1966 年，R. S. Lehman 再次刷新记录，他计算了 250000 (二十五万) 个零点；三年后这一记录又被 J. B. Rosser 改写为 3500000 (三百五十万) ……

Riemann ζ 函数的零点计算步入了快车道！

十四. 最昂贵的葡萄酒

验证了三百五十万个零点虽不足以证明什么， 但对 Riemann 猜想还是有着一定的心理支持作用。不过许多数学家对这点心理支持作用很不以为然，其中有一位数学家最为突出，不仅不以为然，而且还“顶风作案”，跟同事打赌！

这位数学家是德国波恩 Max Planck 数学研究所(Max Planck Institute for Mathematics)的 Don Zagier (1951-)。对 Zagier 来说，区区三百五十万个零点简直就是 zero evidence，因为他认为 Riemann ζ猜想的反例根本就不可能出现在这么前面的零点之中，因此在他看来当时已完成的所有有关零点的计算其实都还远没有涉及到真正有价值的区域。那么要计算多少个零点才可能会对 Riemann 猜想具有判定性的价值呢？Zagier 通过对一些由 Riemann ζ函数衍生出来的辅助函数的研究，认为大约要 300000000 (三亿) 个零点。

Zagier 的怀疑论调很快遇到了对手。二十世纪七十年代初，Max Planck 数学研究所的访客名单中出现了一位铁杆的 Riemann 猜想支持者：Enrico Bombieri (1940-)。 这是一位非同小可的人物，他在不久之后的 1974 年获得了数学最高奖——菲尔兹 (Fields) 奖。Bombieri 深受哲学家 William of Occam 的科学简单性原则(俗称 Occam 剃刀)的影响，对他来说，一个不在 critical line 上的零点就象交响乐中的一个失控的音符，是完全无法令人接受的。

一个怀疑、一个深信，怎么办呢？Zagier 提议打赌。不过人生苦短，两人都意识到自己未必有机会能在有生之年见到 Riemann 猜想被证明或证伪。为了不使赌局太过遥遥无期，双方决定以 Zagier 认为具有判定性价值的前三亿个零点为限。如果 Riemann 猜想在前三亿个零点中出现反例，就算 Zagier 获胜；反之，如果 Riemann 猜想被证明，或者虽然没被证明但在前三亿个零点中没出现反例，则算 Bombieri 获胜 [注三]。 他们定下的赌注为两瓶波尔多葡萄酒 (Bordeaux)。

Zagier 估计这个赌局要分出胜负也许得花上三十年，因为当时计算机的运算能力距离能够计算三亿个零点还相差很远，而且计算 Riemann ζ函数的零点没什么应用价值， 在 CPU 时间十分昂贵的时代并不是人们热衷的计算课题。可是没想到仅仅过了几年，1979 年，由 Richard Brent 领导的一个澳大利亚研究组就把零点计算推进到了前 81000000 (八千一百万)个零点。不久由 Herman te Riele 领导的一个荷兰研究组更是成功地计算出了前两亿个零点。所有这些零点都毫无例外地落在 Riemann 猜想所预言的 critical line 上。 这一系列神速的进展对 Zagier 的钱包显然是大大的凶兆，到这时他已经知道自己大

大低估了计算机领域的发展速度。不过 te Riele 在两亿个零点处终止计算还是让 Zagier 松了一口气，他庆幸道："毫无疑问他们有能力推进到三亿，但感谢上帝，他们没那么做。 现在我总算有几年时间可以喘息了。他们是不会为了多算 50% 而推进的。人们会等待能够算到十亿个零点的那一天，那将是许多年后的事了。"

Zagier 的如意算盘原本打得不错，计算零点不象百米赛跑，在百米赛跑中由于比赛记录已经逼近人类所能达到的速度极限，因此大家不惜为百分之一秒争个你死我活 计算零点却是一条没有尽头的征程，计算能力的发展在相当长的时间内也是没有尽头的。在这种没有尽头的征程上，仅仅多算百分之几十的零点是不够刺激的，人们更感兴趣的是数量级上的推进。这正是 Zagier 认为自己可以喘息几年的心理屏障。可惜人算不如天算。Zagier 万万没有想到他的一位好朋友 Hendrik Lenstra 当时正在荷兰，不仅在荷兰，而且与 te Riele 同在一个城市——阿姆斯特丹！Lenstra 是知道 Zagier 和 Bombieri 的赌局的。如今眼看好戏就要开演了，正自心痒难搔，te Riele 竟然不合时宜地在两亿个零点处停了下来，Lenstra 的那份难受就甭提了(大家以后可得留神好朋友啊！ :-)，套用一句韦爵爷的话说，那真是"生可忍， 熟不可忍"。于是他给 te Riele 做思想工作：你知不知道，如果你算到三亿，Zagier 就会输掉一个赌局！te Riele 一听原来计算零点还有这么伟大的意义，那还等什么？把 Zagier 干掉啊！于是大家一鼓作气把计算推进到了 307000000 (三亿零七百万) 个零点处。那是在 1982 年。

Zagier 输了。

Zagier 兑现了诺言，买来两瓶葡萄酒，Bombieri 当场打开其中一瓶与 Zagier 共享。这一瓶酒，用 Zagier 的话说，是世界上被喝掉的最昂贵的葡萄酒。因为正是为了这两瓶酒，te Riele 特意多计算了一亿个零点。这花费了整整一千个小时的 CPU 时间，而 te Riele 所用的计算机的 CPU 时间在当时大约是七百美元一小时。换句话说，这两瓶酒是用七十万美元的计算经费换来的，被他们喝掉的那一瓶价值达三十五万美元！

喝完了那瓶酒，Zagier 从此对 Riemann 猜想深信不疑。只不过，Bombieri 相信 Riemann 猜想是因为它的美丽，是因为 Occam 剃刀；而 Zagier 相信 Riemann 猜想是因为证据，是因为他觉得证据已经足够强了。

注释

[注一] Hilbert 第十问题是：给定一个具有任意多未知数的 Diophantine 方程，设计一个过程，能用有限多次运算确定该方程是否具有有理整数解。Turing 对计算机及人工智能的研究与此有着密切的关系。

[注二]我有时觉得 Turing 与 John Nash (影片"Beautiful Mind"的主角)颇有相似之处：两人都对纯数学有浓厚的兴趣，研究成果却对应用领域影响深远；两人都对物理学有一定的兴趣；两人都有为军方服务的经历；两人后来的精神世界都偏离了常轨……

[注三]严格讲，他们的条款还忽略了一种可能性，那就是 Riemann 猜想在数学上被证伪，但反例并不在前三亿个零点之中 (或虽然在前三亿个零点之中，但尚未有人进行计算)。显然，这个忽略对 Zagier 比较不利，不过它对赌局后来的发展没有产生影响。

十五. 更高、更快、更强

三亿个零点摆平了 Zagier，但显然远不是对 Riemann ζ 函数非平凡零点进行计算的终点。不过在介绍进一步进展之前我们先要对零点计算做一点补充说明。

当我们说到零点计算的时候，一般人会很自然地认为所谓零点计算，顾名思义就是计算

零点的数值。不知读者在 上一节 时有没有想过这样一个问题：那就是三亿个零点，即使每个只保留十位数字，写下来也有三十亿个数字(如果加上小数点、 等号及零点编号等，则数字还要翻上一番)。以每页三千个数字而论，起码要一百万页纸才能记录下来！当然，计算结果不是非得记录在纸上不可的。但是三十亿个数字差不多是 3GB，这在今天虽然算不了什么，在 1982 年却是非同小可的数量，用任何方式记录都并不容易。以计算机硬盘为例，当时容量为几个 MB 就算很大了，价格十分昂贵，而要想记录三亿个零点却要上千个这样的硬盘！若果真如此，Zagier 岂不还大大低估了他那两瓶葡萄酒的价值？

其实狡猾的 te Riele 并没有计算那些零点的具体数值。事实上除了最初那些小范围的计算外，我们前面介绍的大规模零点计算并不给出零点的具体数值，而只是验证零点是否在 critical line 上。因此，当人们说“计算了前 N 个零点”时，实际指的往往只是验证了前 N 个零点是否位于 critical line 上[注一]。

但是不计算零点的数值，又如何判断零点是否在 critical line 上呢？其实很简单。我们在第十一节中介绍过，要研究 Riemann ζ函数在 critical line 上的零点，只需研究 $Z(t)$ 的符号改变即可。假如在区间 $0<t<T$ 内 $Z(t)$ 的符号改变 N 次，则 Riemann ζ函数在 critical line 上该区间内至少有 N 个零点。另一方面，我们虽不确定是否所有零点都在 critical line 上，却知道它们全部位于 critical strip——$0<\mathrm{Re}(\rho)<1$ 内 (参阅 第七节)，而人们早就知道如何计算 critical strip 内位于区间 $0<\mathrm{Im}(\rho)<T$ 的零点总数 (最早的方法是由 Riemann 本人给出的对 $d\xi(s)/2\pi i\xi(s)$ 沿矩形区域$\{0<\mathrm{Re}(\rho)<1, 0<\mathrm{Im}(\rho)<T\}$ 作边界路径积分—参阅 第五节)。显然，只要我们能够证明：

1.在 critical strip 内位于区间 $0<\mathrm{Im}(\rho)<T$ 的零点总数为 N。

2.在 critical line 上位于区间 $0<t<T$ 的零点至少有 N 个。

就可以推知 Riemann ζ 函数的前 N 个零点全部位于 critical line 上。由于这两者都不涉及零点的具体数值，因此我们可以不计算零点数值就直接证明 Riemann ζ函数的前 N 个零点 (或更一般地，复平面上某个区域内所有的零点) 都位于 critical line 上，这正是大多数零点计算所采用的方法。

对 Riemann ζ函数零点的计算越推进 (即 N 越大)，我们在复平面上沿虚轴方向延伸得就越高 (即 T 越大)。随着计算机运算速度越来越快，te Riele 的三亿个零点的记录很快就失守了。四年后，由他本人及 J. van de Lune 领衔将计算推进到了十五亿个零点。此后 van de Lune 及其他一些人继续进行着零点计算。不过这时已经很少有人象当年的 Turing 那样觉得有可能通过零点计算直接找到 Riemann 猜想的反例，也再没有象 Zagier 那样敢于下注的勇士了。人们在计算零点上的兴趣和投入遂大为下降。这其中一个显著的变化就是逐渐用廉价的小型或微型机取代以往的大型机，且往往使用机器的闲散时间而非正规工作时间来进行计算。尽管如此，计算机技术的神速发展还是抵消了所有这些因素带来的不利影响。零点计算仍在推进着，只是速度变得缓慢起来，这种趋势一直延续到二十世纪末 (2000 年)。

但是到了 2001 年 8 月，德国 Böblingen IBM 实验室的 Sebastian Wedeniwski 启动了一个被称为 ZetaGrid 的计划，建立了迄今为止最强有力的 Riemann ζ函数零点计算系统，重新将零点计算推向了快车道。ZetaGrid 系统将零点计算通过计算机网络分散到大量的计算机上，从而极大地拓展了资源利用面。ZetaGrid 刚启动的时候，加入系统的计算机只有 10 台，半年后就增加到了 500 台，这些都是 IBM 实验室的内部计算机。 一年后，Wedeniwski 将 ZetaGrid 推向了互联网，任何人只要下载安装一个小小的软件包就可以使自己的机器加入 ZetaGrid，此举很快吸引了大量的参与者。如今在 ZetaGrid 上的联网计算机数平均已在一万以上，虽然 ZetaGrid 上的多数计算是利用各台机器的闲散 CPU 时间进行的 (比如通过背景过程或屏保程序) 但由如此大量的计算机所形成的总体运

算能力依然十分可观。截止本文写作之日，ZetaGrid 所计算的零点累计已达 8553 亿个 (其中有六百万个是由本文作者贡献的 :-)，而且还在以大约每天十亿个以上的速度增加着

十六. 零点的统计关联

除了不计算具体数值这一特点外，前面所介绍的那些大规模零点计算还有一个特点，那就是都只针对前 N 个零点。换句话说，所有那些计算都是以第一个零点为起始的。它们所验证都只是复平面上 $0<\mathrm{Im}(\rho)<T$ 之间的零点。除了这类计算外，在零点计算中还有一类计算也十分重要，那就是针对一个虚部很大的区间 $T_1<\mathrm{Im}(\rho)<T_2$ 的计算 (即从某个很大的序号开始的零点计算)。这类计算中最著名的人物是 Andrew M. Odlyzko，他在二十世纪八十年代末和九十年代初对序号在 1020-30769710 和 1020+144818015 间的总计 175587726 个零点进行了计算。2001 年和 2002 年，他更是把计算的起始点推进到了第 10^{22} 和 10^{23} 个零点附近,所计算的零点数目也分别增加到了一百亿和两百亿。Odlyzko 的这些计算不仅所涉及的区域远远超出了 ZetaGrid 的验证范围，而且还包含了对零点数值的计算。这些计算对于研究 Riemann 猜想的意义不仅在于它们提供了有关这一猜想的新的数值证据，更重要的是它们为研究 Riemann ζ 函数非平凡零点在 critical line 上的统计关联提供了数据。这也正是 Odlyzko 进行这类计算的目的。

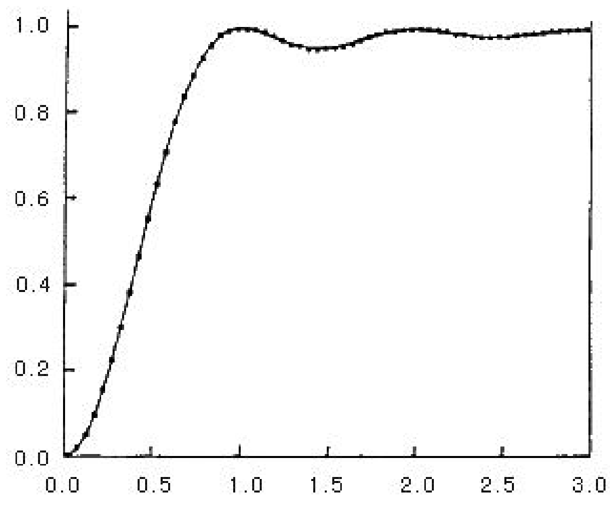
那么 Odlyzko 为什么会研究起零点的统计关联来呢？这还得从二十世纪七十年代初说起。当时英国剑桥大学有位来自美国的研究生叫做 Hugh Montgomery，他所研究的课题是零点在 critical line 上的统计关联。

Montgomery 这个名字不知大家有没有觉得面熟？对了，本系列各篇文章所引的共同题记正是出自此人!

我们以前谈论零点分布的时候，所关心的往往只是零点是否分布在 critical line 上。Montgomery 的研究比这更进一步。他想知道的是，假如 Riemann 猜想成立，即所有零点都分布在 critical line 上，那它们在 critical line 上的具体分布会是什么样的?

在 Montgomery 进行研究的时候虽然已经有 Rosser 对前三百五十万个零点的计算结果 (参阅 第十三节)，但如我们在上文中所说，那些计算并不涉及零点的具体数值，从而无法为他提供统计研究的依据。因此 Montgomery 只能从纯理论的角度来研究零点在 critical line 上的统计关联。

Montgomery 对零点分布的理论研究从某种意义上讲恰好与 Riemann 对素数分布的研究互逆。Riemann 的研究 是着眼于通过零点分布来表示素数分布 (参阅 第五节)，而 Montgomery 的研究则是逆用 Riemann 的结果，着眼于通过素数分布来反推零点分布。



零点的对关联函数

不幸的是，素数分布本身在很大程度上就是一个谜。除 了素数定理外，有关素数分布的多数命题都只是猜测。而素数定理，如我们在 第七节 中看到的，与零点分布的相关性非常弱，不足以反推出 Montgomery 感兴趣的信息。于是 Montgomery 把目光投注到了比素数定理更强的一个命题，那便是 Hardy 与 Littlewood 于 1923 年提出的关于孪生素数分布规律的猜测，即迄今尚未证明的著名的强孪生素数猜想 (有关这一猜想的介绍可参

阅拙作 孪生素数猜想)。Montgomery 以 Riemann 猜想的成立为前提，以 Riemann 的公式及 Hardy 与 Littlewood 所猜测的孪生素数分布规律为依据，研究提出了有关 Riemann ζ 函数非平凡零点在 critical line 上的分布规律的一个重要猜测：

$$\lim_{T\to\infty}\frac{\left|\left\{(t',t'')\mid 0\le t'<t''\le T,\ \frac{2\pi\alpha}{\ln(T/2\pi)}\le t''-t'\le\frac{2\pi\beta}{\ln(T/2\pi)}\right\}\right|}{\frac{T}{2\pi}\ln\frac{T}{2\pi}}=\int_\alpha^\beta\left[1-\left(\frac{\sin\pi t}{\pi t}\right)^2\right]dt$$

上式中 t' 和 t'' 分别表示一对零点的虚部，α 和 β 是两个常数(α < β)。很明显，上式表示的是零点的对关联 (pair correlation) 规律。这一规律被称为 Montgomery 对关联假设 (Montgomery pair correlation conjecture)，其中的密度函数 $\rho(t) = 1-[\sin(\pi t)/\pi t]^2$ 被称为零点的对关联函数 (pair correlation function)。

从上述分布规律中可以看到 lim t→0 ρ(t) = 0，这表明两个零点互相靠近的几率很小。换句话说 Riemann ζ 函数的非平凡零点有一种互相排斥的趋势。这一点与 Montgomery 最初想象的很不相同。Montgomery 曾经以为零点的分布是高度随机的，如果那样的话，对关联函数应该接近于 ρ(t) ≡ 1。这一分布也不同于 Montgomery 当时见过的任何其它统计分布——比如 Poisson 分布或正态分布——中的对关联函数，它与素数本身的分布也大相径庭。这一分布究竟有何深意呢？对 Montgomery 来说还是一个谜。

大家也许还记得，在第五节中我们曾经介绍过 Riemann 提出的三个命题，其中第一个命题(也是迄今唯一被证明的一个)表明在区间 0<Im(ρ)<T 上 Riemann ζ 函数的非平凡零点的数目大约为 (T/2π)ln(T/2π)-(T/2π)。由此不难推知 (请读者自行证明) Riemann ζ 函数相邻零点的间距(即虚部之差) 大约为 Δt~2π/ln(t/2π)。这一间距随 t 而变， 这使得 Montgomery 对关联假设的形式比较复杂。有鉴于此，Montgomery 之后的数学家 (比如 Odlyzko) 对零点的虚部做了处理，引进了归一化的零点虚部：

$$n = (t/2\pi)\ \ln(t/2\pi)$$

利用这一定义，相邻零点的间距被归一化为 Δn~1，而 Montgomery 对关联假设可以简化为 (请读者自行证明)：

$$\lim_{N\to\infty}\frac{\left|\{(n',n'')\mid 0\le n'<n''\le N,\ \alpha\le n''-n'\le\beta\}\right|}{N}=\int_\alpha^\beta\left[1-\left(\frac{\sin\pi t}{\pi t}\right)^2\right]dt$$

Montgomery 对关联假设提出之后，一个很自然的问题就是：零点分布果真符合这一假设吗？这正是 Odlyzko 登场的地方。由于 Montgomery 对关联假设涉及的是对关联在 T→∞ 情形下的极限分布，因此要想对这一假设进行高精度的统计检验，最有效的办法是研究虚部很大的零点的分布，这也正是 Odlyzko 将零点计算推进到 10^{20} 及更高区域的原因。我们在右上方的图中给出了 Montgomery 零点对关联函数 (曲线) 及由 Odlyzko 利用 10^{20} 附近七千万个零点对之进行统计检验的结果 (数据点)。两者的吻合几乎达到了完美的境界。

1972 年春天，刚刚完成上述零点统计关联研究的 Montgomery 带着他的研究成果飞往美国 St. Louis 参加一个解析数论会议。在正式行程之外，他顺道在 Princeton 高等研究所做了短暂的停留。没想到这一停留却在数学与物理间造就了一次奇异的交汇，我们 Riemann 猜想之旅也因此多了一道神奇瑰丽的景致。

注释

[注一] 举个例子来说，虽然早在 1982 年 te Riele 就“计算了”前三亿个零点，但直到

几年后 Odlyzko 与 te Riele 才合伙对区区两千个零点做了真正的数值计算(精度达小数点后一百位)，并以此为基础一举否证了 Mertens 猜想 (参阅第五篇 [注一])。

十七. 茶室邂逅

Montgomery 虽然得到了有关 Riemann ζ函数非平凡零点对关联函数的猜测性结果。但这一结果究竟有何深意，对他来说还是一个谜。他觉得这个结果应该预示着什么东西，可那究竟是什么呢？他并不知道，这多少让他感到有些苦恼。

带着他的研究成果，也带着那几分苦恼，1972 年春天 Montgomery 飞往美国 St. Louis 参加一个解析数论会议。这趟旅行对 Montgomery 有着一举数得的意义。在此前不久他已经接受了一份来自美国 Michigan 大学(University of Michigan) 的工作，因此会后他到 Michigan 大学所在地 Ann Arbor 买了房子(Montgomery 目前仍在 Michigan 大学数学系)。

至此这趟旅行已经获得了精神与物质的双重丰收。但在结束旅程前 Montgomery 还有一事放心不下。

我们在第三节曾经提到 Gauss 有一个“坏毛病”，那就是常常不发表自己的工作，结果使得同时代的许多数学家在研究课题上与他“撞车”(与 Guass 这样的大师玩碰碰车，谁的脑袋先碰破就不必说了)。无独有偶，二十世纪的 Princeton 高等研究所也出了一位有同样“坏毛病” 的数学家，那便是 Atle Selberg (1917-)。Selberg 在 Riemann 猜想的研究中有着极为重要的地位，我们在后文中将会更多地介绍他，这里就不赘述了。让 Montgomery 放心不下的就是自己会不会与 Selberg “撞车”，自己的这项研究工作会不会不幸在 Selberg 的某一叠草稿纸上已经有了？当然，除此之外他也很想听听这位 Riemann 猜想研究中的顶尖高手对自己的这项工作的看法，特别是对结果背后含义的理解。于是在返回英国前他决定在 Princeton 高等研究所做短暂的停留，以便会见一下 Selberg。

Montgomery 如愿见到了 Selberg。 但 Selberg 听完了 Montgomery 的介绍只是礼貌地表示了兴趣，却没有提出具体意见。不过他总算也没有说：“干得不错,小伙子，但是 N 年前我已经证明过这样的结果了”，还是让 Montgomery 松了一口气。

见过了 Selberg，Montgomery 便和朋友 Sarvadaman Chowla (1907-1995) 到 Fuld Hall 去喝下午茶。喝下午茶虽是一种休闲，在 Princeton 高等研究所的学术氛围中却是一个重要的组成部份。在这一时间里来自世界各地、从事不同研究的学者们互相攀谈，交流看法，往往会撞击出一些意想不到的智慧火花。

Montgomery 和 Chowla 正在喝茶闲聊的时候，一位物理学家走了进来。

在 Princeton 高等研究所这样一个科学家阵容豪华得近乎奢侈的地方，随便哪个角落碰上的都可能是非同小可的人物。这位漫步走进茶室的物理学家也不例外。此人在二十世纪中叶曾因证明了量子电动力学的几种形式体系彼此等价而获得了很高的声誉，也为他赢得了 Princeton 高等研究所的终生职位。而这项研究还只不过是他科学生涯中许许多多研究中的一个。他的研究涉及到核物理、凝聚态物理、天体物理，乃至天体生物学等诸多领域。 这位物理学家便是 Freeman Dyson (1923-)。在二十世纪物理殿堂的璀璨群星中 Dyson 当然远不是最杰出的，但那个午后他和 Montgomery 的世界线在高等研究所的短暂交汇却是科学史上一段难忘的佳话，对于 Riemann 猜想的研究来说也是一个奇峰突起的精彩篇章。

Chowla 是一位交际高手，一边和 Montgomery 喝茶聊天，一边仍能眼观六路、耳听八方。Dyson 刚一进门就被他发现了，于是他问 Montgomery:“你见过 Dyson 吗？”，Montgomery 说没有，Chowla 就说我给你引见一下。Montgomery 心想自己做的东西

和 Dyson 八杆子都打不着，再说喝完茶就走人了，何必还特意打扰 Dyson？就说不必了。但Chowla 却是一个从来不把“不”字当成答案的家伙，当下二话不说就把 Montgomery 拽到了 Dyson 跟前 (谢谢 Chowla！)。

就这样 Dyson 和 Montgomery 攀谈了起来。Dyson 问 Montgomery 最近在研究什么？ Montgomery 就把自己对 Riemann ζ函数非平凡零点分布的研究叙述了一下。Dyson 礼貌地听着，他对这一领域并不熟悉。连 Selberg 都没有发表具体的看法，Montgomery 也并不指望这番泛泛介绍会得到比礼貌地点点头更多的回应。

但是当他介绍到自己所猜测的密度函数$\rho(t)=1-[\sin(\pi t)/\pi t]^2$(详见 第十六节) 时，Dyson 的眼睛猛地睁大了！

因为这个让 Montgomery 找不到北，甚至连 Selberg 也看不出端倪来的密度函数对 Dyson 来说却一点也不陌生，那正是随机厄密矩阵(Random Hermitian matrices) 本征值的对关联函数。物理学家们研究这类东西已经有二十年了！

而 Dyson 本人也早在十年前就系统地研究了随机矩阵理论，是这一领域公认的先驱者之一。即使找遍整个世界，也不可能找到一个比 Dyson 更合适的人来和 Montgomery 共喝这杯下午茶了。他们的相遇本身就是一个幸运的奇迹[注一]。

十八. 随机矩阵理论

身为理论物理学家的 Freeman Dyson 如何会研究起随机矩阵理论来的呢？这当然还得从物理学说起。

我们知道在物理学上可以严格求解的问题是少之又少的。而且物理理论越发展，可以严格求解的问题就越少。举个例子来说，在 Newton 引力理论中二体问题可以严格求解，但一般的三体问题就不行[注二]；到了广义相对论中连一般的二体问题也解不出了，只有单体问题还可以严格求解；而到了量子场论中更是连单体问题也解不成了。

另一方面，现实物理中的体系往往既不是单体，也不是二体或三体，而是多体，少则十几、几十 (比如大一点的原子、分子)，多则 10^{23}或更多 (比如宏观体系)。很明显，对现实物理体系的研究离不开各种近似方法。这其中很重要的一类方法就是统计方法，由此形成了物理学的一个重要分支：统计物理。

在统计物理中，人们不再着眼于对物理体系的微观状态进行细致描述(因为这种细致描述不仅无法做到， 而且对于确定体系的宏观行为来说是完全不必要的)，取而代之的是“系综”的概念。所谓“系综”，指的是满足一定宏观约束条件的大量全同体系的集合，这些体系的微观状态具有一定的统计分布，我们感兴趣的体系的宏观状态就由相应物理量的系综平均值所给出。

在传统的统计物理中，组成系综的那些全同体系具有相同的哈密顿量 (Hamiltonian)，只有它们的微观状态才是随机的。但是随着研究的深入，物理学家们开始接触到一些连这种方法也无法处理的物理体系，其中一个典型的例子就是由大量质子中子组成的原子核。这种体系的相互作用具备了所有可以想象得到的 “坏品质” (比如耦合常数很大，不是二体相互作用，不是有心相互作用等)，简直是 “五毒俱全”。对于这种体系，我们甚至连它的哈密顿量是什么都无法确定。 这样的体系该如何处理呢？很显然还是离不开统计的方法。只不过以前在系综中只有各体系的微观状态是随机的，现在却连哈密顿量也不知道了 既然如此，那就只好一不做二不休，干脆把哈密顿量也一并随机化了。由于哈密顿量可以用矩阵来表示，因此这种带有随机哈密顿量的量子统计系综可以用随机矩阵理论来描述。这一点最早是由 Eugene Wigner (1902-1995) 于 1951 年提出的[注三]。

把哈密顿量随机化不等于说对哈密顿量的结构就没有任何限制了。二十世纪六十年代初，Dyson (正是与 Montgomery 在茶室里偶遇的的那位 Dyson) 对随机矩阵理论进行

了深入的研究，并在 1962 年一连发表了五篇非常漂亮的论文。这些论文在随机矩阵理论的发展中具有奠基性的作用。在这些论文中 Dyson 证明了随机矩阵理论可以按照体系在时间反演变换 T 下的性质分为三种类型：

如果体系不具有时间反演不变性，则演化算符为幺正矩阵 (Unitary Matrices)。

如果体系具有时间反演不变性，且 $T^2=I$，则演化算符为正交矩阵(Orthogonal Matrices)。

如果体系具有时间反演不变性，且 $T^2=-I$，则演化算符为辛矩阵(Symplectic Matrices)。

这里 Dyson 用演化算符 U 取代了哈密顿量 H，这两者之间由 $U=\exp(-iHt)$ 相联系。用演化算符的好处是它的参数空间是紧致的。

除了按照对称性对演化算符的结构进行分类外，还有一个需要解决的问题就是哈密顿量的分布函数。Dyson 引进的是 Gauss 型分布，这是数学物理中比较常见的一种分布。在这种分布下具有上述三种对称性的系综分别被称为：Gaussian Unitary Ensemble (GUE), Gaussian Orthogonal Ensemble (GOE)和 Gaussian Symplectic Ensemble (GSE)。

限于篇幅，下面我们就只讨论 Gaussian Unitary Ensemble,它所对应的体系哈密顿量是厄密的，也就是 Dyson 所说的“随机厄密矩阵”。Gaussian Unitary Ensemble 中的随机厄密矩阵的几率测度定义为：

$$P(H)dH = C\exp[-tr(H^2)/2s^2]dH$$

其中 C 为归一化常数，H 为体系的哈密顿量，σ 为标准差，通常取为 $2^{-1/2}$。

对于一个量子体系，能级分布是在理论与观测上都极其重要的性质，这也是随机矩阵理论中物理学家们最感兴趣的东西之一。 物理学家所说的能级用数学术语来说就是哈密顿量的本征值。那么随机厄密矩阵的本征值是怎样分布的呢？分析表明，一个 N 阶随机厄密矩阵的本征值分布密度为：

$$P(\lambda_1, \ldots, \lambda_N)=C\cdot\exp[-\Sigma_i\lambda_i^2]\cdot\Pi_{j>k}(\lambda_j-\lambda_k)^2$$

其中 $\lambda_1, \ldots, \lambda_N$ 为本征值，C 为归一化常数。

通过对这一分布密度的积分，我们可以计算出随机厄密矩阵本征值的各种关联函数。但是这些关联函数的表观复杂程度与本征值的平均间距有很大关系，因此我们要先对本征值做一点处理，以便简化结果。这一处理所依据的是 Wigner 曾经证明过一个结果，那就是当矩阵阶数 $N\to\infty$ 时，n 阶随机厄密矩阵的本征值趋向于区间$[-2(2n)^{1/2}, 2(2n)^{1/2}]$ 上的半圆状分布，即：

$$P(\lambda)\cdot d\lambda=(8n-\lambda^2)^{1/2}/4\pi\cdot d\lambda$$

其中 $P(\lambda)\cdot d\lambda$ 为区间 $(\lambda, \lambda+d\lambda)$ 上的本征值个数。这一规律被称为 Wigner 半圆律 (Wigner Semicircle Law)。利用这一规律，我们对本征值做一个标度变换，引进：

$$\mu = \lambda(8n-\lambda^2)^{1/2}/4\pi$$

可以证明 (请读者自己证明)，这一变换就象我们在第十六节中对 Riemann ζ 函数零点虚部所做的处理那样，将本征值的间距归一化为：$\Delta\mu\sim1$。在这种间距归一化的本征值下关联函数的形式变得相对简单，其中对关联函数的计算结果为：

$$P_2(\mu_1, \mu_2)= 1-[\sin(\pi|\mu_2-\mu_1|)/\pi|\mu_2-\mu_1|]^2$$

看到这里，大家想必也和 Dyson 一样看出来了，随机厄密矩阵本征值的对关联函数正是我们在第十六节中介绍过的，Montgomery 所猜测的 Riemann ζ 函数非平凡零点的对关联函数!当然那时候 Montgomery 用的不是象“对关联函数”这样摩登的术语，事实上“对关联函数”这一术语 Montgomery 在和 Dyson 交谈前连听都没听说过，他自己用的是象“我正在研究零点间距”这样土得掉渣的“白话文”。

在本节快要结束的时候，爱思考的读者可能会提出这样一个问题，那就是为什么要选择 Gauss 型分布？对于这个问题，实用主义的回答是：Gauss 型分布是数学上比较容易处理的 (不要小看这样的理由，当问题复杂到一定程度时这种理由有时是最具压倒性的)；稍为深刻一点的回答则是：Gauss 型分布在固定的$|H|^2$系综平均值下具有最大的熵，换句话说它描述的是在一定约束下具有最大随机性的体系。但是最深刻的回答却是：我们其实并不需要特意选择 Gauss 型分布！ 随机矩阵理论的一个非常引人注目的特点便是：在矩阵阶数 $N\to\infty$ 的极限下它的许多结果具有普适性(即不依赖于哈密顿量的特定分布)。正是这种普适性使得随机矩阵理论在从复杂量子体系的能级分布到无序介质中的波动现象，从神经网络系统到量子混沌，从 $N_c\to\infty$ 的 QCD 到二维量子引力的极为广阔的领域中都得到了应用。

但是把所有这些不同尺度、 不同维度上的应用加在一起，也比不上它与 Riemann ζ 函数非平凡零点分布之间的关联来得神奇。Montgomery 曾经为不知道自己的结果预示着什么而苦恼，现在他知道了那样的结果也出现在由随机矩阵理论所描述的一系列物理现象中。

但是这种存在于象 Riemann ζ 函数非平凡零点分布这样最纯粹的数学与象复杂量子体系、无序介质那样最现实的物理间的神奇关联本身又预示着什么呢？

注释

[注一]有意思的是，在与 Montgomery 的这次 “茶室邂逅”的前一年(即 1972 年)，Dyson 刚写过一篇题为 “Missed Opportunity”(“错过的机会”) 的文章，叙述了科学史上由于数学家与物理学家交流不够而错失发现的一些事例。

[注二]这里“单体”、“二体”、“三体”指的都是点状分布或可视为点状分布的体系。

[注三] 当然，在这一领域中数学家还是要先于物理学家。随机矩阵理论在数学中最早是由 J. Wishart 于 1928 年提出的。