

时域统计分析在疾病传播动力学中的应用*

熊诗杰[†]

(南京大学物理系 南京 210093)

摘要 文章介绍了运用时域统计分析进行疾病传播动力学研究所得到的新结果. 在给定某一类疾病的潜伏期和活动期的时域统计分布参数后, 该疾病在具有一定感染几率的封闭系统中的传播动力学过程即可通过求解概率微分方程来决定. 计算结果表明, 在经过长时间传播后, 该系统的最终被感染率由指数 $\tau_2 = cNt_2$ 决定, 这里 c 是系统内个体的接触感染几率, N 是系统所包含的个体数量, t_2 是疾病活动期的平均时间长度. 如果 $\tau_2 > 3$, 则该系统的最终被感染率可达到 100%, 因此 τ_2 又称为疾病传播的危险指数. 而在疾病传播的初期, 新发病例所出现的振荡和间歇行为则由另一个指数 $\tau_1 = cNt_1$ 决定, 这里 t_1 是平均潜伏期的长度. 这一理论较好地弥补了以通常用的 SIR 模型对时域统计考虑不足的缺陷, 从而对疾病在封闭系统中的传播动力学过程给出理论预测.

关键词 时域统计, 传染病, 传播动力学

Application of time-domain statistical analysis in the dynamics of epidemic outbreaks

XIONG Shi-Jie[†]

(Department of Physics, Nanjing University, Nanjing 210093, China)

Abstract New results in the dynamics of epidemic outbreaks obtained by time-domain statistical analysis are presented. Once the parameters of the time-domain statistical distributions for the incubation period and active period of an epidemic disease are given, the dynamics of the disease outbreak in a closed system with given infection probability can be completely determined by solving stochastic differential-integrative equations. Calculation shows that, after a long enough time, the final infection rate of the system is determined by the index $\tau_2 = cNt_2$, where c is the infection probability of individuals in the system, N the number of individuals, and t_2 the average length of the active period of the disease. If $\tau_2 > 3$, the final infection rate of the system may reach 100%, so τ_2 can also be called the risk index. In the early days of the epidemic, the oscillatory and intermittent behavior of the new-active rate is determined by another index, $\tau_1 = cNt_1$ where t_1 is the average length of the incubation period. This theory overcomes the drawback of the SIR model which does not take sufficient account of the time-domain statistics, and can provide a better theoretical prediction for the infection dynamics in a closed system.

Key words time-domain statistics, epidemic disease, spreading dynamics

众所周知, 统计物理所研究的是由大量微观粒子所组成的宏观系统. 通过对微观粒子运动规律的研究, 可以根据一定的统计法则得出整个宏观体系的各种性质, 如比热、压强、电导、热导等. 近年来, 这

种统计物理的研究方法, 在凝聚态物理和材料物理

* 国家自然科学基金(批准号 60276005)资助项目

2004-08-10 收到初稿, 2004-09-14 修回

[†] 通讯联系人. E-mail: sjxiong@nju.edu.cn

等领域都取得了很大的成功. 一个凝聚态系统, 或一种新材料, 是由大量的原子(包括其中的原子核和核外电子), 以一定的结构组合在一起. 通过对原子核和核外电子微观运动的统计分析, 往往可以成功地预言该系统的各种宏观性质. 在物理学传统领域之外的客观世界中, 也存在各种由大量个体所组成的系统. 整个系统的性质往往也可以由单个个体的行为通过统计分析来确定. 例如, 传染病在一个群体中的传播是由个体之间的接触和传染造成的, 个体的性质和个体的运动规律将决定整个疾病传播过程. 又如, 信息在互联网中的传播是由节点(计算机)之间的连接实现的, 单个计算机的行为和网络连接的结构决定了信息传播的速度和范围. 尽管在真实世界中, 这些由大量的个体所组成的系统是各种各样、千差万别的, 但它们往往会表现出一些共同的统计性质. 从这个意义上说, 统计物理的模型和方法, 在研究这些系统的性质时是十分有用的.

在疾病传播动力学的研究方面, 早在 20 世纪 20 年代, Lowell Reed 和 Wade Hampton Frost 就提出了著名的 SIR 模型. SIR 是英文 susceptible, infective 和 removed 三个单词的缩写, 意即指每个个体在疾病传播过程中可能取的三种状态: 易感的、感染的和被移除的状态. 易感的状态是指还没有被传染, 但也并没有免疫力, 因而是有可能将来被传染的个体. 感染的状态是指已经被感染、正处于生病状态并且可以传染他人的个体, 而被移除的状态是指已经治愈并获得终身免疫力, 或治疗无效而死亡的个体. 在被移除的状态下, 该个体不会再被传染, 也不会传染他人. 每个个体都处在 S, I 和 R 三种状态中的一种, 因此, 三种状态的个体数总和等于该群体总的个体数. 在 SIR 模型中, 每个处于 I 状态下的个体在单位时间(如一天)内都有一定的几率(记作 α)把疾病传染给每个处于 S 状态下的个体, 使其从 S 状态变为 I 状态. 每个处于 I 状态下的个体在单位时间里也有一定的几率(记作 β)被治愈或因治疗无效而死亡, 使其从 I 状态变为 R 状态. 根据几率 α 和 β 的数值, 可以列出微分方程组, 计算出三种状态个体数占总个体数的比率随时间的演变规律^[1]. 由于 SIR 模型没有考虑到被治愈后又重新感染(即没有获得免疫力)的情况, 有人将 SIR 模型改为 SIS 模型, 即只考虑 S 和 I 两种状态, 在感染者被治愈后将回到 S 状态, 而不是进入 R 状态. 当然在 SIS 模型中, 没有考虑到因治疗无效而死亡的情况.

虽然 SIR 模型(或 SIS 模型)已经应用多年, 甚

至被写入流行病学的教科书中, 但在预言疾病传播动力学方面并不理想. 在传染病传播过程中, 很多复杂情况很难用该模型描述. 近年来, 有很多研究者开始提出一些新的理论方案, 试图更准确地描写疾病传播的动力学过程. 这方面的研究引起了极大的兴趣. 这种兴趣源于对艾滋病、SARS 等危害人类健康的严重传染病进行控制的高度迫切性. 这些新的理论方案主要集中在对整个疾病传播网络的空间结构提出更切合实际的模型. 在初始的 SIR 模型中, 没有考虑传播网络的细致结构, 认为处于群体内的某一感染者, 可以以同样的几率传染给该群体内的任一易感染的个体. 但在实际的情况中, 该群体内的不同个体与该感染者的接近程度是不同的, 因此应有不同的被感染几率. 2002 年, Newman 提出应该研究在不同的网络结构中 SIR 模型的动力学行为^[2]. 在这些网络结构中, Watts 和 Strogatz 所提出的小世界网络(small-world network)模型^[3], 引起了广泛的注意^[4-6]. 小世界网络是从固体物理中常用的晶格网络演变而来的. 在晶格网络中, 一个格点(代表一个个体)只能与其最近邻的格点相连接, 这意味着网络中的个体不能自由移动, 代表一种最为固化的网络结构. 在小世界网络中, 晶格网络的部分最近邻连接被打断, 代之以超过最近邻的远程的连接. 这说明每个个体除了可以与其相邻的个体有接触外, 还可以通过运动或其他方式与更远的个体有接触, 当然, 接触的几率随着距离的增加而下降, 这种下降在小世界网络模型中基本上是指指数型的快速下降. 另一种网络是有长程尾连接几率的模型, 称为自由标度网络(scale-free network), 其连接几率随距离的增加呈幂函数型的缓慢衰减, 可以描述在现代社会中, 两个个体虽然相距遥远, 但仍有一定的几率互相接触^[7-9].

虽然小世界网络和自由标度网络等模型充分考虑了疾病传播赖以进行的网络系统的空间结构, 但对个体的感染 - 发病 - 治愈(或死亡)全过程的时域结构, 仍然沿用 SIR(或 SIS)模型中的简单参数 β , 因此对疾病传播动力学中的一些基本特征仍然不能很好地描述. 从疾病控制的观点来看, 以下两个问题是疾病传播动力学必须解决的重要问题:

(1) 疾病长期传播后, 被感染率所能达到的最大饱和值及其所依赖的因素. 一种传染病在一个固定的群体中传播, 被感染的个体在整个群体中所占的比率, 称为被感染率(记作 γ), 它总是随着时间的推移增加的. 一个重要的问题是, 经过足够长的时间

后,被感染率是否会达到 100%? 我们知道在一些非洲国家,艾滋病的被感染率已经达到了一个很高的水平,而且还在缓慢增加,人们非常关心在一段时间后被感染率会否变成 100%,即所有的人都被感染了. 如果情况并没有这么悲观,那么最终的饱和和被感染率会是多少? 通过控制哪些因素可以有效地减少最终的被感染率? 这确实是大家都关心的问题.

(2) 在传染病传播的初始阶段,新增被感染率,即在单位时间里新增的被感染个体在群体中所占的比率(记作 η)是如何随时间变化的? 是单调上升,还是单调下降,或者是振荡型的? 是否会出现间歇现象(即在一段时间内新发病例为零,但之后又开始增加)? 如果出现间歇现象,间歇期的长度与什么因素有关? 对这些规律的了解,有助于在传染病爆发初期分析传播趋势和采取有效的控制措施.

应该说, SIR 模型对以上两个问题都缺乏满意的答案,后来尽管加上对传播网络空间结构的考虑,情况也没有很大的改善. 对第一个问题, SIR 模型给出的最终饱和被感染率是 100%,这似乎与历史上传染病爆发流行的情况不太符合,因为无论在人类或其他哺乳动物中发生的严重传染病流行过程中,都很少出现整体灭绝的情况. 对第二个问题, SIR 模型则很难描述间歇现象,而我们知道,在很多传染病流行过程中,确实是存在间歇现象的.

实际上, SIR 模型中简单的参数 β ,无法很好地描述个体发病的时间过程. 按照 SIR 模型,一个个体从被感染起,在单位时间里就有 β 的几率被治愈(或死亡),也就是说,在被感染的第一时间段他就有被治愈(或死亡)的可能性,而在经过 m 个时间间隔以后,他仍然有 $(1-\beta)^m$ 的几率既没有被治愈也没有死亡. 这种时域结构显然是不合理的,很难想象一个人在被感染后的第一时间就被治愈或死亡,也很难想象一个人在得病后任意长的时间里仍然没有被治愈也没有死亡. 这样就造成了 SIR 模型无法对以上两个问题给出正确的解答.

在文献[10]中,我们提出了个体发病和治愈(或死亡)过程的一个较为合理的时域统计模型:在一个个体被感染后,他要经历一个潜伏期,长度记为 t_1 ,在潜伏期内,他既没有发病,也没有传染性. 经过潜伏期之后,他开始发病,而且产生了传染性,进入了活动期,其长度记作 t_2 ,在这之后,他被治愈或死亡,进入了“被移除”的状态. 这样,原来 SIR 模型中的一个简单参数 β ,被两个时间段的长度 t_1 和 t_2 所代替. 很显然, t_1 和 t_2 与传染病的种类有关,决定了

不同传染病的不同传播动力学行为. 对于同一种传染病而言,由于个体差异,潜伏期和活动期应分布在其平均值附近的一定范围之内,分别在区间 $\left[t_1 - \frac{w_1}{2}, t_1 + \frac{w_1}{2}\right]$ 和区间 $\left[t_2 - \frac{w_2}{2}, t_2 + \frac{w_2}{2}\right]$ 内. 于是在我们的时域统计模型中,总共有四个参数:潜伏期和活动期的平均长度 t_1 和 t_2 ,以及它们在所研究的群体中分布范围的大小 w_1 和 w_2 .

在给定了时域统计模型后,我们来研究这 4 个参数对疾病传播动力学行为的影响. 设在一个有 N 个个体的封闭群体中,从占总个体数的比率为 γ_0 的“种子”(seed)个体开始,爆发了传染病,每个个体的发病过程遵从如上的时域统计模型,而每个被感染的个体在活动期的单位时间内,有 c 的几率可以将疾病传染给群体内任一个尚未受感染的个体. 据此,我们可以列出被感染率 γ 和单位时间里新增病例在群体中所占的比率 η 作为时间的函数所满足的微分方程组^[10]. 在对不同参数值求解了这个微分方程组之后,我们得到了一系列十分有趣且重要的结果.

在图 1 中,我们给出了在 $w_1 = w_2 = 0$ (即所有个体的潜伏期和活动期都是 t_1 和 t_2) 时在不同参数值下 γ 和 η 随时间变化的情况. 图中时间 τ 的单位是 $(cN)^{-1}$. 十分有趣的是,从图 1(a)可以看出,在各种不同的参数下,经过足够长的时间后,被感染率 γ 确实都达到了某一小于 100% 的饱和值,该饱和值主要与一个无量纲的参数 $\tau_2 = cNt_2$ 的数值有关. 这就回答了我们前面提出的第一个问题,即最后这个群体会不会被 100% 感染的问题,与 SIR 模型不同的是,我们对该问题的回答是否定的. 这里我们还找到了决定最终被感染率的重要无量纲参数 τ_2 ,这一点我们下面还要讲到. 而单位时间里新增病例在群体中所占的比率 η 则往往被视为传染病暴发流行的严重程度. 从图 1(b)可以看出,在传染病流行的初期, η 往往会突然上升到一个很大的数值,其后它总体来说是缓慢下降的,但其间会出现振荡,甚至出现间歇现象,即在某一时间段内新发病例为零,其后又开始上升. 可以说,这些曲线可以很好地重现历史上许多传染病流行的动力学行为,特别值得注意的是,在 τ_2 给定后,振荡周期和间歇行为主要取决于另一个无量纲参数 $\tau_1 = cNt_1$,亦即由该种传染病的潜伏期长度来决定. 这样,我们也很好地回答了前面的第二个问题.

为了进一步研究到底是什么因素影响了最终的

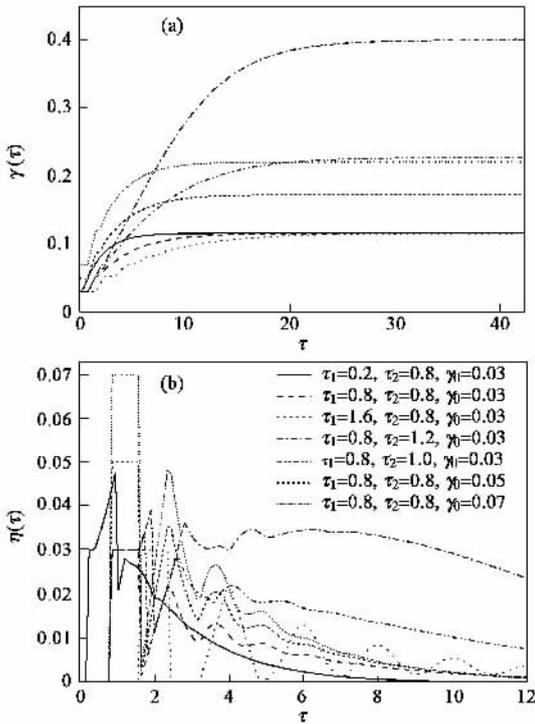


图1 (a)被感染率 γ 随时间的变化行为 (b)单位时间里新增病例在群体中所占的比率 η 随时间的变化行为[时间 τ 以 $(cN)^{-1}$ 为单位. 各条曲线所使用的参 τ_1 , τ_2 和 γ_0 在图 (b) 中给出]

被感染率 $\gamma(\infty)$ 的,我们在图2中给出了 $\gamma(\infty)$ 与最初的种子比率 γ_0 以及前述的无量纲参数 τ_2 的关系.从图中可以清楚地看出, $\gamma(\infty)$ 几乎与最初的种子比率 γ_0 无关,而主要由无量纲参数 τ_2 决定.在 $3 \geq \tau_2 \geq 0.5$ 的范围内, $\gamma(\infty)$ 随 τ_2 急剧单调增加,其数值还小于 100%; 在 $\tau_2 > 3$ 时, $\gamma(\infty)$ 的数值已接近于 100%. 于是,我们可以把无量纲参数 τ_2 定义为传染病传播的一个危险指数,如果该危险指数小于 3,传染病还不至于遍及整个群体,而如果该危险指数大于 3,基本上可认为传染病最终可在该群体中 100% 传播.这应该说是我们的时域统计模型对传染病传播动力学的预后情况的一个极为重要的结论.由于 $\tau_2 = cNt_2$, 这里 c 是群体内接触传染的几率, t_2 是活动期的长度,而 N 是该群体的大小.从传染病的种类来看,活动期 t_2 越长,以及传染几率 c 越大的传染病越危险.因此,艾滋病虽然传染几率较小,但其活动期很长,而 SARS 虽然活动期很短,但其传染几率很大,二者都是很危险的传染病.除了疾病本身的因素以外,群体的大小 N 是另一个决定性因素.在一个小的群体里,疾病最终可能不会 100% 传播,而同样的事情如果发生在一个大的群体里,却有可能 100% 传播.这初看起来好像不可思议,但我们可以

以注意一下如下的例子.相信禽流感病毒在很早以前就有了,但在家家户户各自养鸡的时代,一个村庄不过几十只、上百只鸡,即使发生了禽流感,最终也只有百分之几,也就是几只鸡死亡,根本不会引起社会关注.而现在进入了大规模养鸡的时代,一个鸡场动辄几万、几十万只鸡,那么一旦发生禽流感,如上所述,由于 N 很大,会使危险指数 τ_2 超过 3,那么最终将有近 100% 的鸡,也就是几万或几十万只鸡感染死亡,必然引起社会震动.实际上,在长期的生物进化过程中,许多小的物种往往比大的物种(如恐龙)有更强的生命力,这当然有很多其他的原因,但面对传染病传播,小的物种由于 N 小而比大的物种有更小的危险指数,应该也是原因之一.从疾病控制的角度来看,如何极大地降低危险指数就是关键所在.以下几种方法应该都是有效的方法 (1) 提高治疗效率,使病人及早康复以减少活动期的长度 t_2 ; (2) 切断传播途径以减少传播几率 c ; (3) 把一个大的群体分割成很多互相隔离的小群体以减少 N ,这在疾病暴发流行的危险时期不失为极有效的应急方法,在流行期学校紧急停课等方法应属于这一类.

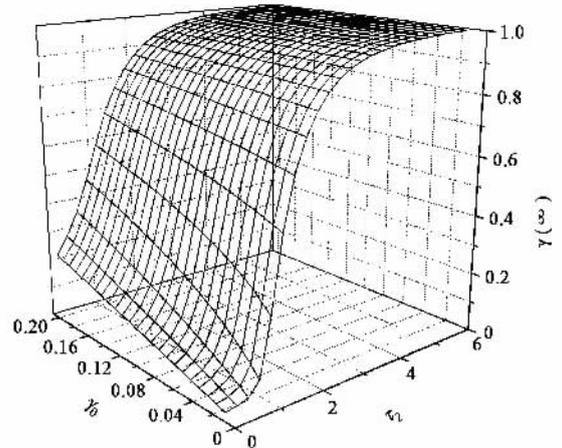


图2 最终被感染率 $\gamma(\infty)$ 与最初的种子比率 γ_0 以及无量纲参数 τ_2 的关系

以上的结论是在潜伏期和活动在群体中的分布长度 w_1 和 w_2 为零的情况下得出的.如果 w_1 和 w_2 不为零,计算表明,最终的被感染率 $\gamma(\infty)$ 并不受到影响,前面的结论仍然成立,只有最初的振荡行为受到影响.在图3中,我们给出了在不同的 w_1 和 w_2 的值下 η 和 γ 随时间演变的情况.从图中我们可以看到,在 w_1 和 w_2 增大的情况下,长时间的饱和行为没有受到影响,只有短时的振荡振幅减少了,原有的间歇区减少甚至消失.

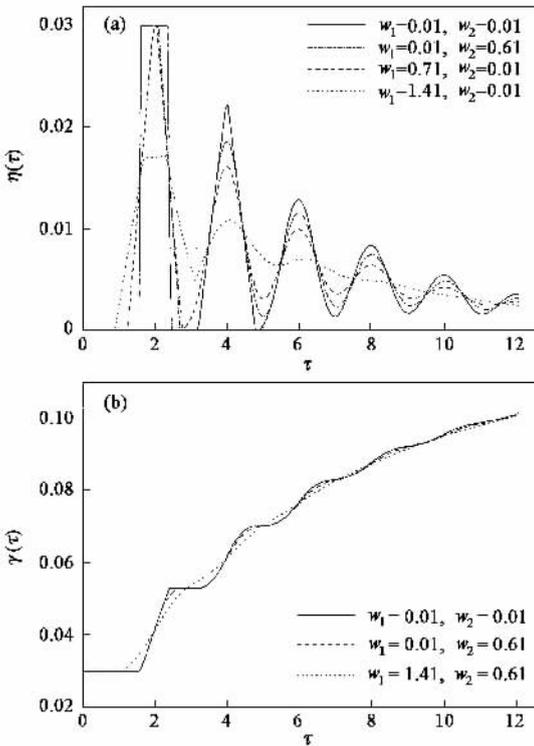


图3 (a)单位时间里新增病例在群体中所占的比率 η 随时间的变化 (b)被感染率 γ 随时间的变化[时间 τ 以 $(cN)^{-1}$ 为单位.其余参数是 $\tau_1=1.6, \tau_2=0.8, \gamma_0=0.03$]

从以上的分析可以看出,对应于潜伏期和活动期长度的无量纲参数 τ_1 和 τ_2 是决定疾病传播动力学行为的最重要的参数.在文献[10]中,我们给出了有关计算的一系列详细的方程和公式,由此不但可以对传播的长期行为进行描述,也可以对传播的短期行为进行描述,在很大程度上克服了广泛应用的 SIR 模型的缺陷,而得到与实际相符合的结论.

参考文献

- [1] Bailey N T J. The Mathematical Theory of Infectious Diseases and its Applications. New York: Hafner Press, 1974
- [2] Newman M E J. Phys. Rev. E 2002 66 016128
- [3] Watts D J, Strogatz S H. Nature (London), 1998, 393: 440
- [4] Moore C, Newman M E J. Phys. Rev. E, 2000, 61: 5678
- [5] Strogatz S H. Nature (London), 2001, 410: 268
- [6] Zhu C P, Xiong S J, Tian Y J et al. Phys. Rev. Lett. 2004, 92: 218702
- [7] Barabási A -L, Albert R. Science, 1999, 286: 509
- [8] Barabási A -L, Albert R, Jeong H. Physica A, 1999, 272: 173
- [9] Pastor-Satorras R, Vespignani A. Phys. Rev. E, 2001, 63: 066117
- [10] Xiong S J. Phys. Rev. E, 2004, 69: 66102

· 物理新闻和动态 ·

超流固体

在今年1月的Nature杂志上,发表了美国宾洲大学的 Moses Chan 和 Eun-Seong Kim 两位教授的一个有关超流固体的有趣实验.具体的实验工作如下:将一块微小的圆盘悬挂在细棒间,盘内充满了多孔性的玻璃材料——维克玻璃,在玻璃内注入氦-4原子,然后将整个样品冷冻.当样品温度达到 2K,压力为 63 大气压时,氦-4原子转变为固态.让这个在海绵体内具有固态氦-4原子的小盘产生运动,这时小盘慢慢地发生像钟摆一样的振动,同时记录下它的共振频率.如果将小盘的温度继续下降,当温度达到 175mK 时,会发生一种相变,即氦-4仍保持着固体状态,但其行为却类似于超流体.

为了更好地观察这个现象,Chan 和 Kim 将小盘的力学性能作一定的改变,这时振动也会随着改变.当达到某一特定温度时,固体的转动惯量会发生一个突变,它的行为类似于超流体.这时可形象地看到氦原子能无摩擦地在多孔介质(维克玻璃)内滑动.在科学上应该如何设想这样一种固体在另一种固体中随意滑动的物理图像呢? Chan 和 Kim 教授利用量子理论来解释他们的实验.他们认为氦-4原子在低温下能无摩擦地运动是由于处于固态的氦-4原子在量子测不准原理的作用下还保持着一定的运动能量.对于较轻的原子,如氢原子,它的零点能相对来说是比较大的,同时维克玻璃内又存在着许多空穴,这就为氦原子在空穴内穿梭往返提供了可能性.氦-4原子晶体是被一个单一的波函数所控制,它的行为类似于在一个单量子态上发生了玻色-爱因斯坦凝结,这实际上是一个量子涨落现象.

为了进一步地研究这个现象,在9月份的Science Express杂志上,两位教授对论证“超流”特性又发表了他们新的实验结果.这次他们在作实验时,不再让氦-4原子放在维克玻璃体内,而是安置在一个不封闭的环形管道内,管道可以在一个容器内自由转动.第二步仍然是对样品降温、加压,从而使氦-4原子固化,这时从特征振动的性质中可以确定氦-4原子处于固态,而不是液态.随着温度的继续下降,当温度达到 230mK 以下时,转动开始发生突然的变化.根据这新的实验结果,Chan 和 Kim 教授认为固体氦内有 1.5% 的原子发生了变形,它们可以自由流动.所以这时发生的超流物理过程确实是处在固态氦的情况下,而不是处在液态氦的状态.因此我们可以说,这是一个无摩擦的超流固体.

(云中客 摘自 Nature, 15 January 2004; Science Express, 3 September 2004)