忙总谈数学：微积分

January 22, 2022

Abstract

这一篇帖子主要介绍人类如何从一个基于几何直观或直觉的计算技巧或计算方法，进化到逻辑基础严密的公理体系的例子，想说明人类抽象的另外一个方向：语言抽象（结构抽象已经在介绍伽罗华群论时介绍过）。

现代数学基础有三大分支：分析，代数和几何。这篇帖子以尽量通俗的白话介绍数学分析。数学分析是现代数学的第一座高峰。为了让非数学专业的人能够看下去，采用了大量描述性语言，所以严谨是谈不上的，只能算瞎扯。

最后为了说明在数学中，证明解的存在性比如何计算解本身要重要得多，用了两个理论经济学中著名的存在性定理（阿罗的一般均衡存在性定理和阿罗的公平不可能存在定理）为例子来说明数学家认识世界和理解问题的思维方式，以及存在性的重要性：阿罗的一般均衡存在性，奠定了整个微观经济学的逻辑基础–微观经济学因此成为科学而不是幻想或民科；阿罗的公平不可能存在定理，摧毁了西方经济学界上百年努力发展，并是整个应用经济学三大支柱之一的福利经济学的逻辑基础，使其一切理论成果和政策结论成为泡影。

豆瓣小组“管理实践与学习”
https://www.douban.com/group/542139/

瞎扯数学分析1、微积分
来自: wxmang 2017-02-04 13:42:16
https://www.douban.com/group/topic/96542437/

这一篇帖子主要介绍人类如何从一个基于几何直观或直觉的计算技巧或计算方法，进化到逻辑基础严密的公理体系的例子，想说明人类抽象的另外一个方向：语言抽象（结构抽象已经在介绍伽罗华群论时介绍过）。

为了让非数学专业的人能够看下去，采用了大量描述性语言，所以严谨是谈不上的，只能算瞎扯。

现代数学基础有三大分支：分析，代数和几何。这篇帖子以尽量通俗的白话介绍数学分析。数学分析是现代数学的第一座高峰。

最后为了说明在数学中，证明解的存在性比如何计算解本身要重要得多，用了两个理论经济学中著名的存在性定理（阿罗的一般均衡存在性定理和阿罗的公平不可能存在定理）为例子来说明数学家认识世界和理解问题的思维方式，以及存在性的重要性：阿罗的一般均衡存在性，奠定了整个微观经济学的逻辑基础–微观经济学因此成为科学而不是幻想或民科；阿罗的公平不可能存在定理，摧毁了西方经济学界上百年努力发展，并是整个应用经济学三大支柱之一的福利经济学的逻辑基础，使其一切理论成果和政策结论成为泡影。

一、微积分

数学分析是微积分基础上发展起来的，所以先说说微积分。

微积分的基本思想是以直为曲，也即用直线来逼近曲线，在中国古代，刘徽，祖冲之计算圆周率用的割圆术就是典型的微积分方法，三国时期的刘徽在他的割圆术中提到的“割之弥细，所失弥小，割之又割，以至于不可割，则与圆周和体而无所失矣。”魏晋南北朝时期的祖冲之说的更简单：以曲为直逼近。在古代巴比伦，希腊都用这种方法来处理曲线计算问题，有史可查的记录是公元前三世纪，古希腊的阿基米德计算抛物弓形的面积、球和球冠面积、螺线下面积和旋转双曲体的体积时，就用了直线逼近。

所以在牛顿（Newton）和莱布尼茨（Leibniz）发明微积分之前，很多实际上的微积分的工具已经开始运用在科学和工程之中。例如法国的费尔玛、笛卡尔、罗伯瓦、笛沙格；英国的巴罗、瓦里士；德国的开普勒；意大利的卡瓦列利等人都用这种以直为曲的逼近方法计算工程问题。

但是微积分为什么说是十七世纪牛顿和莱布尼茨发明的呢，我觉得主要是两点：第一点是引入了函数概念来描绘变量；第二点是发明了一套符号体系，可以计算各种初等函数微分（初等函数简单说就是多项式函数、幂函数、指数函数、对数函数、三角函数、反三角函数，以及由这些函数经过有限次四则运算或函数的复合而得的所有函数）。

牛顿和莱布尼茨发明的最原始的微积分可以解决以下问题：

求即时速度的问题；求曲线的切线；求函数的最大值和最小值；求曲线长、曲线围成的面积、曲面围成的体积、物体的重心、一个体积相当大的物体作用于另一物体上的引力等等。

牛顿和莱布尼兹最本质的贡献是把求切线问题（微分学的中心问题）和求积问题（积分学的中心问题）变成一个问题。这就是著名的牛顿–莱布尼兹公式。

牛顿和莱布尼茨建立微积分的基本思想是以曲为直，逐步逼近，其中创造是引入了无穷小量Δ，因此微积分也称为无穷小分析。

不过他们两个有区别，牛顿从运动角度入手，莱布尼茨从几何角度路入手。

牛顿在1671年写了《流数法和无穷级数》，这本书直到1736年才出版，它在这本书里指出，变量是由点、线、面的连续运动产生的，否定了以前自己认为的变量是无穷小元素的静止集合。他把连续变量叫做流动量，把这些流动量的

导数叫做流数。牛顿在流数术中所提出的中心问题是：已知连续运动的路径，求给定时刻的速度（微分法）；已知运动的速度求给定时间内经过的路程（积分法）。

莱布尼茨1684年发表世界上最早的微积分文章：《一种求极大极小和切线的新方法，它也适用于分式和无理量，以及这种新方法的奇妙类型的计算》，创立了现代的微分符号和基本微分法则（远远优于牛顿的符号，现在使用的微积分通用符号就是当时莱布尼茨创造的），1686年，莱布尼茨发表了人类第一篇积分学的文章。

微积分的创立，极大地推动了数学的发展，过去很多初等数学束手无策的问题，运用微积分，往往迎刃而解。例如牛顿应用微积分及微分方程从万有引力定律推导出了开普勒行星运动三定律。

微积分也极大的推动天文学、力学、物理学、化学、生物学、工程学等的发展。

由于争抢微积分发明权，欧洲大陆的数学家和英国数学家的长期对立，英国数学陷入牛顿的“流数术”中停步不前，英国数学后来比欧洲整整落后了一百年。

虽然原始微积分是一种强大计算工具，但是从逻辑上讲，牛顿和莱布尼茨的工作都是很不完善的，他们为了计算微分，引入的在无穷和无穷小量概念，其实没有说清楚是个什么东西，例如牛顿的无穷小量，有时候是零，有时候不是零而是有限的小量；莱布尼茨干脆回避解释。无穷小的逻辑基础存在的问题导致了第二次数学危机的产生（这个在介绍现代数学基础的帖子里已经介绍了，不重复）。

19世纪初，法国的柯西对微积分的理论进行了认真研究，建立了极限理论，后来德国的魏尔斯特拉斯进一步的严格化，使极限理论成为了微积分的坚定基础。才使微积分在逻辑上站住脚，而不仅仅是一种计算工具。

微积分的基础概念是函数和极限。前者是微积分的工作对象，后者是微积分的基本工作技巧。

1 函数

函数概念是人类一个很伟大的发现，价值不下于对于数的发现，也是高度抽象的产物。

不过函数的思想却很早，至少在公元前就有了：因果关系，也即有因必有果，一个因对应一个或多个果，或者一个果对应多个因。

这在中国《易经》中已经有成熟的体现（其实《易经》就是64变量的函数论），正因为有了这种因果关系概念，中国远古时代我们先人就有了成熟精妙的辩证法（比黑格尔的辩证法高级多了，精细多了）。西方辩证法也是在有了成熟的函数概念后才成熟的。恩格斯就说过：“数学中的转折点是笛卡儿的变数，有了变数，运动进入了数学；有了变数，辩证法进入了数学”。

不过近代函数概念直接来源于代数方程中对不定方程的求解。

笛卡儿在1637年出版的《几何学》中，引入了现代函数的思想。英国人格雷果里在1667年论文《论圆和双曲线的求积》给出了函数的定义：从一些其他量经过一系列代数运算或任何其他可以想象的运算而得到的一个量。这里的运算指的是加减乘除开方五种代数运算以及求极限运算。

不过现在我们看到的函数定义来自于德国人莱布尼兹，他在1673年论文中，把任何一个随着曲线上的点变动而变动的几何量，如切线、法线、点的纵坐标都称为函数；并且强调这条曲线是由一个方程式给出的。直接定义了：函数表示依赖于一个变量的量。

紧接着函数概念被不断改进，第一个重要改进是瑞士人约翰.伯努利于1698年给出的：由变量和常量用任何方式构成的量都可以叫做的函数。这里的任何方式包括了代数式和超越式。

第二个重要改进是1748年欧拉在《无穷小分析引论》中给出的函数定义：变量的函数是一个解析表达式，它是由这个变量和一些常量以任何方式组成的。现代函数的符号就是欧拉发明的。欧拉还区分了显函数和隐函数、单值函数和多值函数、一元函数和多元函数等。

1775年，欧拉在《微分学》一书中，给出了函数的另一定义：如果某些变量，以这样一种方式依赖于另一些变量，即当后者变化时，前者也随之变化，则称前面的变量为后面变量的函数。这个定义，为辩证法数学化打开了大门。

第三次重要改进是从函数的几何特性开始的，是1746年达朗贝尔给出的，把曲线称为函数（因为解析表达式在几何上表示为曲线）。但是后来欧拉发现有些曲线不一定是由单个解析式给出的，因此提出了一个新的定义：平面上随手画出来的曲线所表示的x与y的关系。即把函数定义为由单个解析式表达出的连续函数，也包括由若干个解析式表达出的不连续函数（不连续函数的名称是由欧拉提出的）。

在整个十八世纪，函数定义本质就是一个解析表达式（有限或无限）。

第四次最重要的改进是1821年柯西在《解析教程》中，给出了如下函数定义：在某些变量间存在着一定的关系，当一经给定其中某一变量的值，其他变量的值也随之确定，则将最初的变量称为自变量，其他各个变量称为函数。这个定义把函数概念与曲线、连续、解析式等纠缠不清的关系给予了澄清，也避免了数学意义欠严格的变化一词。函数是用一个式子或多个式子表示，甚至是否通过式子表示都无关要紧。

不过函数精确定义是德国人狄利克里于1837年给出的：若对x（$a \le x \le b$）的每一个值，y总有完全确定的值与之对应，不管建立起这种对应的法则的方式如何，都称y是x的函数。这一定义彻底地抛弃了前面一些定义中解析式的束缚，强调和突出函数概念的本质，即对应思想。

对应思想是人类伟大的发现，后来的映射，同构，同态等等概念来源于此，这是这个概念最伟大的地方。

当然我们知道狄利克里伟大，主要不是他给出函数的科学定义，而是他给出了著名的狄利克里函数，这个函数是难以用简单的包含自变量x的解析式表达的，但按照上述定义的确是一个函数。

为使函数概念适用范围更加广泛，人们对函数定义作了如下补充："函数$y = f(x)$的自变量，可以不必取[a,b]中的一切值，而可以仅取其任一部分"，换句话说就是x的取值可以是任意数集，这个集合中可以有有限个数、也可以有无限多个数，可以是连续的、也可以是离散的。这样就使函数成了一个非常广泛的概念。但是，自变量及函数仍然仅限于数的范围，而且也没有意识到"函数"应当指对应法则本身。

最后，我们要说说现代数学理解的函数（来自于美国人维布伦）：设集合X、Y，如果X中每一个元素x都有Y中唯一确定的元素y与之对应，那么我们就把此对应叫做从集合X到集合Y的映射，记作$f : X \to Y$，$y = f(x)$。

不过从布尔巴基以后，基于数学结构的函数概念更进一步抽象，从函数、映射进化到关系：

1939年布尔巴基用集合之间的关系定义了函数：设E和F是两个集合，E中的每一个元素x和F中的每一个元素y之间的一个关系f称为函数，如果对每一个$x \in E$，都存在唯一的$y \in F$，它们满足给定的关系。记作$fE \to F$。在布尔巴基的定义中，E和F不一定是数的集合，函数是集合之间的一个关系。也即设集合E和F，定义E与F的积集$E * F$如下：$E * F = \{(x, y)|x \in E, y \in Y\}$。积集E*F中的一个子集$f$称为E与F 的一个关系，若$(x, y) \in f$，则称$x$ 与y有关系f，记为xfy,若(x, y)不属于f，则称x与y无关系f。设f是x与y的关系，即$f \in X * Y$，如果$(x, y) \in f, (x, z) \in f$，必有$y = z$，那么称$f$ 为X到Y 的映射或函数。

这个定义回避了对应这种模糊不清的描述语言，而且把函数从单纯的数的概念推广到一切对象，例如结构，图像，集合等等。

不过微积分要处理的函数概念还是原始的，甚至只能处理初等函数。特点就是函数自变量的变化范围是数域，也即函数定义域与因变量的变化范围值域都是数域。这就是微积分的工作对象。这个对象可以描述一部分基于初等函数规律描述的变量跟结果的因果关系，通过对这种因果关系的分析和计算，人类就能预测或控制符合相应初等函数规律描述的事件或事物的因果关系，例如各种工程设备，武器系统等等，就能建立工业文明。

2 极限

极限是微积分的主要工作技巧。整个数学分析就是建立在极限概念上（包括级数）来处理初等函数因果关系的一门学科。

极限技巧一般是：对无法把握的连续变量，用可以计算的序列（例如数列，时间序列，多项式序列等等）逐步逼近变量，并能够证明这些序列可以无限逼近所求的未知量，然后计算这个序列的极限就可得到变量。

极限思想是微积分的基本思想，函数的连续性，导数以及定积分等等都是借助于极限来定义的。

所以可以说：数学分析就是用极限思想来研究函数的一门学科。

极限的思想在刘徽割圆术就有了，但是仅仅是一种计算方法，而不是一个思维方式。真正的现代极限思想来自于16世纪荷兰人斯泰文计算三角形重心过程中，用逐步逼近方式逼近重心。

牛顿和莱布尼茨最早并不是用极限思想来建立微积分的，他们的概念基础是无穷小，但是由于无穷小是个逻辑上有瑕疵的概念，导致微积分的逻辑基础无法自洽。例如牛顿用路程的改变量ΔS与时间的改变量Δt之比$\Delta S/\Delta t$ 表示运动物体的平均速度，让Δt 无穷小，得到物体的瞬时速度，并由此引出导数概念和微分，他并没有极限概念，他说：“两个量和量之比，如果在有限时间内不断趋于相等，且在这一时间终止前互相靠近，使得其差小于任意给定的差，则最终就成为相等”。这是一种几何直观而不是逻辑，就像小孩在纸上顺便划一下圆，就说是太阳。所以牛顿说不清楚他理解的无穷小到底是是什么。其实牛顿的说法如果用极限概念，很容易在逻辑上说清楚：如果当变量（例如时间t)无限增大或变量的差无限接近0时（$\Delta t \to 0$），则$\Delta S/\Delta t$无限地接近于常数A，那么就说$\Delta S/\Delta t$ 以A为极限，这个极限就是s（路径函数）在t_0时的导数。

不过上述无限的概念仍然是几何直观的，并没有用逻辑描述出无限这个过程是什么，也没有定量地给出ΔS和Δt两个无限过程之间的数量联系，所以在逻辑

上仍然有漏洞。

所以牛顿和莱布尼兹的微积分不断收到怀疑和攻击，例如最常见的质疑是贝克莱大主教的：在瞬时速度概念中，究竟Δt是否等于零？如果说是零，怎么能用它去作除法呢？如果它不是零，又怎么能把包含着它的那些项去掉呢？这就是数学史上所说的无穷小悖论。

牛顿由于没有极限概念，无法回答这种质疑，只能混战。主要原因是微积分起源于人类计算需要从常量扩展到变量，但是牛顿采用处理常量的传统思想来处理变量。

18世纪，罗宾斯、达朗贝尔与罗依里埃等人明确表示极限是微积分严格化的基础。其中最接近现代定义的是达朗贝尔的极限定义：一个量是另一个量的极限，假如第二个量比任意给定的值更为接近第一个量。但是这些定义都无法摆脱对几何直观的依赖。例如什么叫“接近”，逻辑上的含义是什么，其实还是几何直观。

现代极限概念来自于柯西，19世纪，柯西出版的《分析教程》定义：当一个变量逐次所取的值无限趋于一个定值，最终使变量的值和该定值之差要多小就多小，这个定值就叫做所有其他值的极限值，特别地，当一个变量的数值（绝对值）无限地减小使之收敛到极限0，就说这个变量成为无穷小。

柯西把无穷小视为以0为极限的变量，也即无穷小不是似零非零，无穷小非零，只是其极限为零。

魏尔斯特拉斯把柯西的语言翻译成$\varepsilon-\delta$语言，给微积分提供了严格的理论基础。所谓$\lim a_n(n\to\infty)=A$，是指：如果对任何$\varepsilon>0$，总存在自然数N，使得当$>N$时，不等式$|a_n-A|<\varepsilon$恒成立。

这个定义，借助不等式而不是几何直观，通过ε和N之间的关系，定量刻划了两个无限过程之间的联系。这个定义中，涉及到的仅仅是数及其大小关系，此外只是给定、存在、任取等词语，已经摆脱了“趋近”一词，不再求助于运动的直观。

这个定义，本质揭示了无限与有限有本质的不同：无限个数的和不是一般的代数和，它是部分和的极限，是动态过程，而非静态计算结果。举例来讲，用任何静态计算，都无法计算出变速直线运动的瞬时速度，因为速度是变量。这其实就是量变和质变的一个例子：量变能引起质变。例如对任何一个圆内接正多边形来说，当它边数加倍后，得到的还是内接正多边形，是量变而不是质变；但是，不断地让边数加倍，经过无限过程之后，多边形就变成圆，多边形面积便转化为圆面积，这就是量变到质变，这就是极限概念的本质。极限是区分初等数学和高等数学的分界线，初等数学处理静态问题，高等数学可以处理非静态问题了，例如求瞬时速度、曲线弧长、曲边形面积、曲面体体积等问题。

极限概念中，最重要的定理，非魏尔斯特拉斯的多项式逼近连续函数定理莫属，这个定理的简单表述是：闭区间上的连续函数可由多项式一致逼近。

这个定理意味着任何连续函数，都能构造一个多项式函数来逼近它，而多项式函数的导数，微分，积分的计算，简单易行，也即这个定理解决了连续函数的近似计算的逻辑基础问题：存在性。

这个定理最著名的证明是苏联数学家伯恩斯坦构造的著名的伯恩斯坦多项式，这个方法开启了函数构造法这一研究领域（当然对周期性的函数，还可以用三角级数，也即傅利叶级数逼近）。用多项式函数或三角级数逼近连续函数，是现代工程解决问题的主要方法，例如通信领域，如果不懂傅利叶级数，基本寸步难行，在流体力学、结构力学和弹性力学领域，不用多项式函数逼

近，也基本无法计算海量的变量函数。函数构造方法其实是计算数学算法的基础（伯恩斯坦多项式符号太多，无法介绍，有兴趣可以上网搜索：伯恩斯坦多项式即可，有魏尔斯特拉斯定理用伯恩斯坦多项式证明的全过程）。

魏尔斯特拉斯本人最初的证明，是使用的核函数（正态核），并将核函数展开成一致收敛的幂级数，截取前面有限部分就构造出了逼近多项式。现在教材上选取的核函数是Landau核，这个核函数本身就是多项式，因此相比原证明减少了一步，但本质没有改变。魏尔斯特拉斯本人最初的证明不如伯恩斯坦的证明那么直截了当，那么优美（可以翻教科书参考，如果想详细了解过程，可以看菲赫金哥尔茨的《微积分学教程》，这是经典微积分教材）。当然这个定理最直观的证明是勒贝格的折线逼近法：闭区间上的连续函数可以用折线逼近（可以查书）。

极限是微积分的核心概念，微积分处理初等函数变化，一般都涉及无穷概念，无穷概念只有从极限角度理解，才能正确描述和把握，其实描述极限的语言体系是$\varepsilon-\delta$语言是一个相当于公理体系的定义，$\varepsilon-\delta$意义下的极限是一种公理定义下的逼近，这种逼近不是几何描述的，所以没有逻辑悖论的可能。

逼近的常见技巧是放缩和夹逼，也即不等式是极限的主要技巧。

微积分中讨论的连续函数、导数、定积分、级数的敛散性、多元函数的偏导数，广义积分的敛散性、重积分和曲线积分与曲面积分等等概念都是基于极限的思想方法给出。

3 连续

前面说过，微积分主要对象是初等函数，初等函数的本质性质就是连续，就像一元n次方程的根的本直性质的是对称一样，这是很本质的核心问题，当然微积分必须抓住。

所以换句话说，微积分主要工作对象就是连续函数。其实人类在直到牛顿莱布尼兹时代，并不知道还有非连续的函数概念。预先假定都是连续的，而且他们对连续函数理解仅仅是几何直观，把能一笔画成的曲线所对应的函数叫做连续函数。例如伽利略所研究的落体运动，开普勒所研究的绕日运转的行星所扫描的扇形面积，牛顿所研究的流等都是连续变化的量。

所谓连续，直观解释就是运动变化的过程连绵不断，连续函数就是刻画变量连续变化的数学模型。

微积分是以直为曲的，所以对连续函数也要进行这种处理，例如柯西和魏尔斯特拉斯就用离散的多项式来逼近连续函数，这就是极限理论的由来，有了极限，才开始真的能够把握连续函数的性质。

最早人类理解连续函数，就是当x逐渐改变时，函数$f(x)$的相应变动也是逐渐的，不会有任何突增或突减的跳跃式振荡。但这种理解毫无用处，因为既不能计算，也不能控制。

函数连续的精确表述：设函数$f(x)$在点x_0 的某一去心邻域内有定义，任给ε大于零，存在δ 大于零，当$|x-x_0|<\delta$ 时，有$|f(x)-f(x_0)|<\varepsilon$,则称函数$f(x)$在$x_0$点连续。

这就是数学分析的基本语言：$\varepsilon-\delta$语言，不熟悉这套语言体系，无法学会数学分析。

用$\varepsilon-\delta$语言定义的连续函数，就能计算其极限问题，这是微积分的重要内容，因为微分本质就是计算极限。

而连续函数求极限这种复杂问题本质是可以转化为求函数值的问题的，这就可以大大简化求极限难度。

我们知道，函数的连续性是一个局部性质，对区间也不例外。但如果是闭区间上的连续函数，却能把局部性质转化为整体性质，像闭区间上连续函数的有界性、最大最小值性、介值性、根的存在性、一致连续性等。

用$\varepsilon-\delta$语言，我们就能把握连续函数的性质：

连续函数的局部性质：若函数f 在点x_0 连续，则f 在点x_0 有极限，且极限值等于函数值$f(x_0)$。根据这个性质，可以容易证明下述定理：

局部有界性定理：若函数f 在点x_0 连续，则f 在x_0 的某邻域$U(x_0)$内有界。

局部保号定理：若函数f 在点x_0 连续，且$f(x_0) > 0$（或< 0），则对任何正数$r < f(x_0)$（或$r < -f(x_0)$），存在某$U(x_0)$，使得对一切$x \in U(x_0)$有$r < f(x)$（或$r < -f(x)$）。

四则运算定理：若函数f和g在点x_0 连续，则$f \pm g$，$f * g$，f/g（这里$g(x_0) \neq 0$）也都在点x_0 连续。

复合函数定理：若函数f 在点x_0 连续，g在点u_0连续，$u_0 = f(x_0)$,则

$\lim g(f(x))(x \to x_0) = g(\lim f(x))(x \to x_0) = g(f(x_0))$

海涅（Heine)定理：$\lim f(x)(x \to x_0)$ 存在的充分必要条件是对任给的序列$\{x_n\}$，若满足$\lim x_n(n \to \infty) = x_0(x_n \neq x_0)$，则有$\lim f(x_n)(n \to \infty)$ 存在。

最大、最小值定理：若函数f 在闭区间$[a, b]$上连续，则f 在$[a, b]$上有最大值与最小值；或称函数f 在$[a, b]$上达到最大值。

推论（有界性定理）：若函数f 在闭区间$[a, b]$上连续，则f 在$[a, b]$上有界。

介值性定理：设函数f 在闭区间$[a, b]$上连续，且$f(a) \neq f(b)$。若μ为介于$f(a)$与$f(b)$之间的任何实数（$f(a) < \mu < f(b)$或$f(a) > \mu > f(b)$），则至少存在一点$x_0 \in (a, b)$使得$f(x_0) = \mu$。

根的存在定理：若函数f 在闭区间$[a, b]$上连续，且$f(a)$与$f(b)$异号，则至少存在一点$x_0 \in (a, b)$使得$f(x_0) = 0$。即方程$f(x) = 0$在(a, b)内至少有一个根。

反函数连续定理：若函数f 在$[a, b]$上严格单调并连续，则反函数f^{-1}在其定义域$[f(a), f(b)]$或$[f(b), f(a)]$上连续。

初等函数的连续定理：任何初等函数在它的定义域上都连续。

4 导数

导数最初定义是1823年柯西在《无穷小分析概论》中定义的：如果函数$y = f(x)$在变量x的两个给定的界限之间保持连续，并且我们为这样的变量指定一个包含在这两个不同界限之间的值，那么是使变量得到一个无穷小增量。

现在导数定义是19世纪60年代魏尔斯特拉斯用$\varepsilon-\delta$语言定义的：设函数$y = f(x)$ 在点x_0 的某个邻域内有定义，当自变量x在x_0 处有增量Δx，$(x_0 + \Delta x)$也在该邻域内时，相应地函数增量$\Delta y = f(x_0 + \Delta x) - f(x_0)$,如果任意给$\varepsilon > 0$, 存在常数$a$和$\delta > 0$,当$|\Delta x| < \delta$时,使$|\Delta y/\Delta x - a| < \varepsilon$，则称函数$y = f(x)$ 在点x_0 处可导，并称这个极限为函数$y = f(x)$ 在点x_0 处的导数,记为$f'(x_0)$，也记作$y'|_{x=x_0}$或 $\mathrm{d}y/\mathrm{d}x|_{x=x_0}$。

导数的几何直观就是函数形成的曲线在一点的切线的斜率。

最早导数主要用于求变速运动的瞬时速度（计算弹头的穿透能力或动能必须知道弹头接触目标的瞬时速度）和求曲线上一点的切线。牛顿从第一个问题出发，莱布尼兹从第二个问题出发，分别给出了导数的概念。

牛顿的想法很直观，如一辆汽车在10小时内走了600公里，它的平均速度是60公里/小时。但在实际行驶过程中，是有快慢变化的，不都是60公里/小时。设汽车所在位置s与时间t的关系为：$s = f(t)$,那么汽车在由时刻t_0变到t_1这段时间内的平均速度是：

$(f(t_1) - f(t_0))/(t_1 - t_0)$，当$t_1$与$t_0$无限趋近于零时，汽车行驶的快慢变化就不会很大，瞬时速度就近似等于平均速度。

自然就把当$t_1 \to t_0$时的极限$\lim(f(t_1) - f(t_0))/(t_1 - t_0)$作为汽车在时刻$t_0$的瞬时速度，这显然就是导数。

显然根据上述定义，导数是通过极限对函数进行局部的线性逼近，所以导数是函数的局部性质。一个函数在某一点的导数描述了这个函数在这一点附近的变化率。

显然不是所有的函数都有导数（例如产生突变点，奇点的函数就没有导数），一个函数也不一定在所有的点上都有导数。

若某函数在某一点导数存在，则称其在这一点可导，否则称为不可导。显然很容易证明：可导的函数一定连续；不连续的函数一定不可导。

如果函数$y = f(x)$ 在开区间内每一点都可导，就称函数$f(x)$在区间内可导。这时函数$y = f(x)$ 对于区间内的每一个确定的x值，都对应着一个确定的导数，这就构成一个新的函数，这个函数为原来函数$y = f(x)$ 的导函数，记作y'、$f'(x)$、 $\mathrm{d}y/\mathrm{d}x$或 $\mathrm{d}f(x)/\mathrm{d}x$，简称导数。

显然，导数运算满足一下性质：

$$(u \pm v)' = u' \pm v';\ (uv)' = u' * v + u * v';\ (u/v)' = (u' * v - u * v')/v^2$$

根据上导数定义和性质，很容易计算出一些常见函数的导数：

$y = x^n$，$y' = n * x^{n-1}$；
$y = a^{bx}$，$y' = b * a^{bx} * \ln a$；
$y = a^u$，$y' = u' * a^u * \ln a$；
$y = e^{bx}$，$y' = b * e^{bx}$；
$y = e^u$，$y' = u'e^u$；
$y = \log_a x$，$y' = 1/(x \ln a)$；
$y = \ln x$，$y' = 1/x$；
$y = \sin x$，$y' = \cos x$；
$y = \cos x$，$y' = -\sin x$；
$y = \tan x$，$y' = \sec^2(x)$；
$y = \cot x$，$y' = -\csc^2(x)$；
$y = \sec x$，$y' = \sec x * \tan x$；
$y = \csc x$，$y' = -\csc x * \cot x$；
$y = \arcsin x$，$y' = 1/(1 - x^2)^{1/2}$；
$y = \arccos x$，$y' = -1/(1 - x^2)^{1/2}$；
$y = \arctan x$，$y' = 1/(1 + x^2)$；
$y = \mathrm{arccot} x$，$y' = -1/(1 + x^2)$；
$y = \sinh x$，$y' = \cosh x$。

在实际上应用中，大部分常见的函数都上述函数的和、差、积、商或相互复合的结果。所以一般情况下，函数的导函数计算是简单容易的。

导数的几个用途：

判别单调性：若导数大于零，则单调递增；若导数小于零，则单调递减；导数等于零为函数驻点，不一定为极值点。

求极值：如果存在一点，使得导数在之前区间上都大于等于零，而在之后区间上都小于等于零，那么是一个极大值点，反之则为极小值点。

自然推论：若已知函数为递增函数，则导数大于等于零；若已知函数为递减函数，则导数小于等于零。

判断函数凹凸性：如果函数的导函数在某个区间上单调递增，那么这个区间上函数是向下凹的，反之则是向上凸的。如果二阶导函数存在，如果在某个区间上二阶导数恒大于零，则这个区间上函数是向下凹的，反之这个区间上函数是向上凸的。曲线的凹凸分界点称为曲线的拐点。

导数的最著名应用是中值定理和洛必达法则。

中值定理应包括罗尔中值定理、拉格朗日中值定理、柯西中值定理、泰勒中值定理。

罗尔中值定理：如果函数$f(x)$满足：在闭区间$[a,b]$上连续；在开区间(a,b)内可导；在区间端点处的函数值相等，即$f(a)=f(b)$，那么在(a,b)内至少有一点ξ $(a<\xi<b)$，使得$f'(\xi)=0$。

几何上，罗尔定理含义是一条连续的曲线弧，如果除端点外处处有不垂直于x轴的切线，且两端点的纵坐标相等，则弧上至少有一点的切线是水平的。

拉格朗日定理：如果函数$f(x)$满足：在闭区间$[a,b]$上连续；在开区间(a,b)内可导，那么在(a,b)内至少有一点ξ $(a<\xi<b)$，使等式$f(b)-f(a)=f'(\xi)(b-a)$成立。

柯西定理：如果函数$f(x)$及$F(x)$满足：在闭区间$[a,b]$上连续；在开区间(a,b)内可导；对任一$x\in(a,b)$，$F'(x)\neq 0$，那么在(a,b)内至少有一点ξ，使等式$[f(b)-f(a)]/[F(b)-F(a)]=f'(\xi)/F'(\xi)$成立。

泰勒公式：若函数$f(x)$在开区间(a,b)有直到$n+1$阶的导数，则当函数在此区间内时，可以展开为一个关于$(x-x_0)$多项式和一个余项的和：

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2}(x-x_0)^2+\frac{f'''(x_0)}{3!}(x-x_0)^3+\ldots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+R_n$$

其中$R_n=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$，这里$\xi$ 在x和x_0之间，该余项称为拉格朗日型的余项。（$f^{(n)}(x_0)$是$f(x_0)$的n阶导数，不是$f(n)$与x_0 的相乘）

推论：麦克劳林公式：

若函数$f(x)$在开区间(a,b)有直到$n+1$阶的导数，则当函数在此区间内时，可以展开为一个关于x多项式和一个余项的和：

$$f(x)=f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\frac{f'''(0)}{3!}x^3+,,+\frac{f^{(n)}(0)}{n!}x^n+R_n$$

其中$R_n=\frac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1}$,这里$0<\theta<1$。

达布定理：若函数$f(x)$在$[a,b]$上可导,则$f'(x)$在$[a,b]$上可取$f'(a)$和$f'(b)$之间任何值。

推广：若$f(x),g(x)$均在$[a,b]$上可导,并且在$[a,b]$上,$g'(x)\neq 0$,则$f'(x)/g'(x)$可以取$f'(a)/g'(a)$与$f'(b)/g'(b)$之间任何值。

洛必达法则：设当$x\to a$时，函数$f(x)$及$F(x)$都趋于零；在点a的去心邻域内，$f'(x)$及$F'(x)$都存在且$F'(x)\neq 0$；当$x\to a$时$\lim f'(x)/F'(x)$存在(或为无穷大)，那么$x\to a$ 时$\lim f(x)/F(x)=\lim f'(x)/F'(x)$。

又设当$x \to \infty$时，函数$f(x)$及$F(x)$都趋于零；当$|x| > N$时$f'(x)$及$F'(x)$都存在，且$F'(x) \neq 0$；当$x \to \infty$时$\lim f'(x)/F'(x)$存在(或为无穷大)，那么$x \to \infty$时$\lim f(x)/F(x) = \lim f'(x)/F'(x)$。

中值定理经常用于证明方程根的存在性，证明恒等式，证明不等式，研究函数的单调性，求函数极限（用罗必达法则求0/0，∞/∞函数极限是常用手段），求函数的极值与最值，讨论函数的凸凹性，求函数的拐点，求函数的渐近线，描绘函数的图象等等。具体例子可以查教科书。

5 微分

其实导数和微分概念是一致的，没什么更多可说的。

函数$y = f(x)$的微分$\mathrm{d}y = f'(x)\,\mathrm{d}x$。可导与可微是等价的，若求出了函数在一点的导数，再乘以$\mathrm{d}x$即得该点的微分；若求出了函数在一点的微分，再除以$\mathrm{d}x$即得该点的导数；因此导数又叫做微商。

需要注意的是函数在x点的微分是自变量增量的线性函数，因为微分是对函数的局部变化的一种线性描述。如果一个非线性函数某点可微，其在某点的自变量有一个微小的改变时，函数的变化可以分解为两个部分。一个部分是线性部分：在一维情况下，它正比于自变量的变化量△x，可以表示成△x和一个与△x无关，只与函数及有关的量的乘积；在更广泛的情况下，它是一个线性映射作用在△x上的值。另一部分是比△x更高阶的无穷小，也就是说除以△x后仍然会趋于零。当改变量很小时，第二部分可以忽略不计，函数的变化量约等于第一部分，也就是函数在x处的微分。

所以微分主要用于计算函数值的近似值。

但是不是所有的函数的变化量都可以分为以上提到的两个部分。若函数在某一点不可微，就无法用线性函数逼近。

在现代微积分中，微分被定义为将自变量的改变量映射到变化量的线性部分的线性映射。这个映射也被称为切映射。给定的函数在一点的微分如果存在，就一定是唯一的。

微分有以下运算法则：

连锁律：$\mathrm{d}y/\mathrm{d}x = \mathrm{d}y/dz * dz/\mathrm{d}x$;

乘法律：$\mathrm{d}(uv)/\mathrm{d}x = u(\mathrm{d}v/\mathrm{d}x) + v(\mathrm{d}u/\mathrm{d}x)$;

除法律：$\mathrm{d}(u/v)/\mathrm{d}x = (v(\mathrm{d}u/\mathrm{d}x) - u(\mathrm{d}v/\mathrm{d}x))/v^2$

$\mathrm{d}y/\mathrm{d}x$被称为一阶导数，$\mathrm{d}(\mathrm{d}y/\mathrm{d}x)/\mathrm{d}x = \mathrm{d}^2y/\mathrm{d}x^2$被称为二阶导数，以此类推，$\mathrm{d}^ny/\mathrm{d}x^n$被称为$n$阶导数。

稍微多说一句是法线。曲线上一点的法线和那一点的切线互相垂直，微分可以求出切线的斜率，自然也可以求出法线的斜率。函数$y = f(x)$在(x_0, y_0)点切线的斜率为$m = \mathrm{d}y/\mathrm{d}x$在$(x_0, y_0)$的值，那么法线的斜率为$-1/m$。

6 积分

积分原始思想的萌芽很早，甚至早于微分思想，主要用于计算物体运动的路程、变力作功以及由曲线围成的面积和由曲面围成的体积等问题，现在资料据说古希腊德莫克利特、阿基米德、中国的刘徽都用积分思想计算过面积和体积，当然这些方法都建立在特殊的技巧之上，不具有一般性，也没有逻辑基础

保证其是正确的。再晚一点，开普勒的“同维无穷小方法”、卡瓦列利的“不可分量法”、费马的“分割求和方法”更是典型的积分思想。

不过真正的积分发明者还是牛顿与莱布尼兹，因为他们揭示了微分与积分的内在联系–微积分基本定理，也即牛顿–莱布尼茨公式。

积分是微积分的一个核心概念。通常分为定积分和不定积分两种。

定积分严格的数学定义是黎曼用的方式极限给出的，把曲边梯形设想为一系列矩形组合的极限，也即对于一在区间[a,b]上之给定非负函数$f(x)$，$f(x)$所代表的曲线与Ox坐标轴所夹图形的面积

$$S = \int_b^a f(x)\,\mathrm{d}x = \lim \sum f(t_i)(x_{i+1} - x_i) \quad (i = 0, \ldots, n-1,\ n \to \infty)$$

上述符号定义：在闭区间$[a,b]$中取一个有限的点列$a = x_0 < x_1 < x_2 < \ldots < x_n = b$。每个闭区间$[x_i, x_{i+1}]$构成一个子区间。定义$\lambda$为这些子区间长度的最大值：$\lambda = \max(x_{i+1}, x_i)$，其中$0 \le i \le n-1$。

一个闭区间$[a,b]$进行分割$a = x_0 < x_1 < x_2 < \ldots < x_n = b$后，在每一个子区间中$[x_i, x_{i+1}]$取出一点$x_i \le t_i \le x_{i+1}$。

黎曼定义的积分的就是是微分的无限积累，或者说定积分是无限个无穷小量之和。核心思想就是试图通过无限逼近来确定这个积分值。

所以定积分是一种极限,这种极限不同于数列的极限也不同于函数的极限。它是和式的极限，对于体现自变过程的变量的每一个值，不仅区间的分法有无穷多种，而且对于每一个分法，介点也有无穷多种取，因而相应的和式一般有无穷多个值。但它仍然有着与数列极限、函数极限的本质上的相同之处，即当$[x_i, x_{i+1}]$无限变小时，相应的一切和式与某一定数的距离能够变得并保持任意的小。

微积分的最初发展中,定积分即黎曼积分。在实变函数中,可以利用测度论将黎曼积分推广到更加一般的情况，如勒贝格积分。

显然，黎曼积分定义有一个自然问题就是这个黎曼和式是不是一定有极限，极限与子区间划分方法有无关系。

前者就是所谓的可积问题，后者是极限收敛问题。

决定是否可积一般依赖于四个因素：函数、区间、区间的分法、介值的的取法。

很容易证明，当函数在区间上可积时，，不依赖于区间的分法与介值的取法，函数积分数字只与函数和区间两个因素有关。所以在可积的条件下，当求某函数在指定区间上的定积分时，往往可以取一个特殊的分法（如n等分），取介值为划分内的特殊点（如左或右端点）。

可以证明下述结论：

★可积函数必有界，有界函数不一定可积，无界函数一定不可积；

★连续函数一定可积；

★有有限个间断点的有界函数一定可积；

★有无限多个不连续点的单调函数一定可积；

★区间上有无限个不连续点的有界函数（只要间断点的测度为0）也可积。

定积分的主要应用是求和，例如平面图形的面积，求已知截面面积的立体的体积，求旋转体的体积，求曲线的弧长，求旋转曲面的面积，求变力所作的功，计算运动物体的路程，以及物体之间的万有引力等等。另外，定积分可以作为定义函数的一种新的工具，例如连续函数的变上限积分是函数的一个原函

数，又知道某些函数的原函数并不是初等函数。如椭圆积分就不是初等函数，这时我们就把这个积分本身，作为此函数的定义，此为出发点来研究函数。

微积分最基础的定理是牛顿和莱布尼茨分别独自发现的：

一个可积函数在区间$[a,b]$上的定积分等于它的任意一个原函数在区间$[a,b]$上的增量。也即：如果函数$f(x)$在区间[a,b]上可级，并且存在原函数$F(x)$，则：

$$\int_b^a f(x)\,\mathrm{d}x = F(b) - F(a)$$

这个发现给定积分提供了一个有效而简便的计算方法，大大简化了定积分的计算过程。

这个定理是微积分存在的基础，但是证明极其简单。

证明：因为函数$f(x)$在区间[a,b]上可积，任取区间的分割$a = x_0 < x_1 < x_2 ... < x_n = b$，在区间$[x_i, x_{i+1}](i = 1, 2...n-1)$上任取一点$\zeta_i$,则有$\lim \sum f(\zeta_i)(x_{i+1} - x_i)(i = 1, \ldots, n-1;\ n \to \infty) = \int_b^a f(x)\,\mathrm{d}x$，

由于函数f是函数F的导函数，所以根据拉格朗日中值定理得$F(x_i) - F(x_{i-1}) = f(\zeta_i)(x_{i+1} - x_i)$ 其中$\zeta_i \in (x_{i-1}, x_i)$,因此有

$$F(b) - F(a) = \lim \sum F(x_i) - F(x_{i-1})(i = 1, \ldots n-1;\ n \to \infty)$$

$$= \lim \sum f(\zeta_i)(x_{i+1} - x_i)(i = 1, \ldots n-1;\ n \to \infty) = \int_b^a f(x)\,\mathrm{d}x$$

牛顿-莱布尼茨公式简化了定积分的计算，只要知道被积函数的原函数，总可以求出定积分的精确值或一定精度的近似值。

牛顿-莱布尼茨公式是联系微分学与积分学的桥梁，它是微积分中最基本的公式，它证明了微分与积分是可逆运算，同时在理论上标志着微积分完整体系的形成，从此微积分成为一门真正的学科。

利用牛顿一莱布尼茨公式可以证明定积分换元公式，积分第一中值定理和积分型余项的泰勒公式。牛顿-莱布尼茨公式还可以推广到二重积分与曲线积分，从一维推广到多维。牛顿-莱布尼茨公式促进了其他数学分支的发展，在微分方程，傅里叶变换，概率论，复变函数等数学分支中都要用到。

下面说说不定积分。不定积分是是已知导数求原函数，用公式表示是

$$\int f'(x)\,\mathrm{d}x = f(x) + c$$

而前面已经说了，定积分是求面积（Riemann和的极限），不定积分只是求导数的逆运算，所以不定积分与定积分是完全不同的两个概念。但是，牛顿莱布尼兹公式把它们连接在一起。

不过，函数在所讨论区间上的Riemann和的极限的存在性不取决于该函数的不定积分的存在性，函数在所讨论区间上的不定积分的存在性也不取决于该函数的Riemann和的极限的存在性。

因为容易证明：

函数可积不一定该函数存在原函数：因为$f(x)$在区间[a,b]上连续，在区间[a,b]上有界，且只有有限个第一类间断点，和在区间[a,b]上单调有界,则$f(x)$都在在$[a,b]$上可积，由牛顿莱布尼兹公式知道，一个函数如果可导，那么它的导

函数是不可能存在第一类间断点的，所以说一个函数如果存在第一类间断点，那么它是不会有原函数的，也即可积并不能保证有原函数。

函数连续只是可导的必要条件，而非充分条件（如果一个函数可导，其必然连续。如果一个函数连续，则不一定可导，如$Y = |X|$)。

同时，也容易证明，函数有原函数但该函数不一定可积。例如，函数$y = x^{3/2}\sin(1/x)$各点可导，但由于在闭区间[-1,1]上有无界点，故在[-1,1]上不可积。

所以函数可积问题，是传统微积分没能解决的一个问题（有些函数是连续的但处处不可微，有的函数的有限导数并不黎曼可积，一个黎曼可积的函数列收敛到的那个函数不一定是黎曼可积的等等），直到实变函数发展起来，扩展了可积的概念，例如勒贝格积分，也扩展了基于勒贝格测度理解的连续函数的概念，这个问题才圆满解决。

显然，因为牛顿-莱布尼兹微积分基本公式，导数的公式逆向就是初等函数的积分公式，不必多说。

积分计算有非常多的技巧，换元，变量替换，逼近，因式分解等等，可以看教科书，里面有非常多的计算技巧例子。多做习题。华罗庚是世界现代数学家中计算能力名列前茅的变态，他的很多发现或定理证明都是算出来的，晚年的华罗庚为保证自己思维状态，每天没事干就是算积分玩，而且是极难的积分。这个不是传说，是亲眼所见。原来的科大数学系学生（77,78,79三级），计算积分和矩阵，能力在中国所有大学中，无人能及，科大学生不能把华罗庚的线性代数的打洞公式和积分的变换技巧用得风生水起，都不算合格学生。

7 多元微积分

传统多元微积分的基本概念都是一元微分与积分的基本推广，1687年牛顿就提出了偏导数和重积分的思想，欧拉在1769年给出了二重积分及其累次积分与换元计算方法，拉格朗日在1773年给出了三重积分及其累次积分与换元计算方法，雅可比在1833年给出了变量替换中的雅可比矩阵表达。不过当自变量是多元变量时，导数的概念已经不适用了（尽管可以定义对某个分量的偏导数），但仍然有微分的概念。

下面我们只介绍把一元函数微积分二元函数微积分情况，因为扩展到多元函数是类似的。

7.1 二元函数连续性的定义

多元函数微积分的推广，最初是从几何角度开始的。

二元函数$u = f(x, y)$的变量(x, y)在一个平面直角坐标系中代表一个动点P，它的全部可能的位置形成一个平面点集S。从而函数关系f便把动点P的每一个位置(x, y)对应到变量u的一个惟一确定的数值（函数值）$f(x, y) = f(P)$。于是整个函数便表现为变量u按照这个对应关系随着动点P在定义域S上变化而变化，这样，二元函数的概念便同一元函数的一致。

当动点P由一个位置$P(x, y)$变到另一个位置$P_1(x_1, y_1)$时,这变化由它的位移向量$\overrightarrow{PP_1} = \{x_1 - x, y_1 - y\} = \{\Delta x, \Delta y\} = \Delta P$来刻画,这变化的大小便由这向量的长度$|\Delta P| = ((\Delta x)^2 + (\Delta y)^2)^{1/2}$ 来度量。相应的u 的变化Δu其大小由$|\Delta u|$来度量。

$$\Delta u = f(p_1) - f(p) = f(x + \Delta x, y + \Delta y) = f(x, y)$$

于是多元函数在一点P 处的连续性也可以用一元函数连续型定义，也即，在P_1无限趋近于P的过程中，$|\Delta u|$随着$|\Delta P|$而无限变小。这就是说，多元函数u连续，就是任意给定$\varepsilon > 0$，都存在一个$\delta > 0$，使得只要$|\Delta P| < \delta$，就有$|\Delta u| < \varepsilon$。

所以多元连续函数的基本性质也同一元连续函数的一样：

★多元函数在一有界闭集S上定义，其在S上处处连续，则至少在某一点处达到最小值m,又至少在某一点处达到最大值M；

★多元函数连续性在整个集合S上是一致的（即δ 不依赖于P而对于S上的每个点P都有效）；

★如果S是连通的（即S上每两点都能够用完全位于S上的一条折线连接起来）, 则每一个中间值$\mu(m \leq \mu \leq M)$都是某一点处的函数值；

★多元函数如果连续，它在S的每个内点处都可以分解成一元的情形：函数u在一点P的某个邻域(δ) 内处处连续,则必定在其内部的一个方邻域$[\delta]$上一致连续，而在这个方邻域上的变化量具有向量分解式：$\Delta u = \Delta_x u + \Delta_y u$。式中$\Delta_x u = f(x + \Delta x, y) - f(x, y)$,$\Delta_y u = f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y)$

分别作为一元函数$g(x) = f(x, y)$，$h(y) = f(x + \Delta x, y)$，显然其连续性分别关于$y$或$x + \Delta x$是一致的（即相应的$\delta$ 不依赖于y或$x + \Delta x$ ）。

7.2 偏导数

定义了多元函数连续性，就能定义导数了。显然用$\Delta u = \Delta_x u + \Delta_y u$和$g(x) = f(x, y)$，$h(y) = (x + \Delta x, y)$能够证明在$|\Delta P|$趋向0的过程中，变化量$\Delta u$随$\Delta x$、$\Delta y$ 趋向0 的依赖关系。

这就要用到一元函数g, h变化率，即导数$g(x)$、$h(y)$。假定g, h的导数它们在$P(x, y)$的附近都存在，并分别记为$f(x, y), f(x + \Delta x, y)$,

$$\partial u/\partial x = f'_x(x, y) = \lim(f(x + \Delta x, y) - f(xy))/\Delta x(\Delta x \to 0)$$

$$\partial u/\partial y = f'_y(x, y) = \lim(f(x, y + \Delta y) - f(xy))/\Delta x(\Delta x \to 0)$$

这种对自变量之一（其余作为参变量）的导数称为偏导数。利用这些偏导数的存在和一元微分学的中值定理，可以得到：

$$\Delta u = \Delta_x u + \Delta_y u = f'_x(x, y)\Delta x + f'_y(x + \Delta x, y + \theta\Delta y)\Delta y + \alpha(\Delta x)$$

式中θ 介于0到1之间,α 为无限小量。当偏导数连续时，可以进一步写成：

$$\Delta u = \partial u/\partial x \Delta x + \partial u/\partial y \Delta y + \alpha(\Delta x) + \beta(\Delta y)$$

α、β 为无限小量。

7.3 全微分

函数u在点P处是可微的定义：

$$\Delta u = \Delta_x u + \Delta_y u = f'_x(x, y)\Delta x + f'_y(x + \Delta x, y + \theta\Delta y)\Delta y + \alpha(\Delta x)$$

表明，在点P处，变化量Δu随着Δx、Δy趋向0的过程中，存在着近似线性的依赖关系：

$$\Delta u = A\Delta x + B\Delta y + \alpha\Delta x + \beta\Delta y$$

,

式中主要部分的系数A、B不依赖于Δx、Δy，而余项部分的系数α、β是无限小量。

并把这个线性主要部分为u的一个(全)微分，记为

$$\mathrm{d}u = A\Delta x + B\Delta y$$

令$\Delta x \to 0, \Delta y = 0$或$\Delta x = 0, \Delta y \to 0$，即可推出:$A = \partial u/\partial x, B = \partial u/\partial y,$

所以只要微分存在，它的系数就必然是偏导数，因而是惟一的。

在某些特殊情形,这些偏导数都存在，$\mathrm{d}u = \partial u/\partial x\Delta x + \partial u/\partial y\Delta y$关系却不成立；所以不同于一元函数的情形：只有偏导数的存在还不能保证微分存在。

不过偏导数的连续性可以保证微分存在。也即函数是连续可微的，所以这时u的微分可以写成$\mathrm{d}u = \partial u/\partial x\,\mathrm{d}x + \partial u/\partial y\,\mathrm{d}y$。具体证明可以查教科书，这里不啰嗦，因为很简单（因为x,y是动点P的连续函数）。

7.4 变量替换

变量x, y既然当作动点P的函数，也就可以表达为：动点P在任一别的坐标系(r, s)中的坐标的函数：$x = \varphi(r, s), y = \psi(r, s)$

假定这些坐标函数也在其定义域S上是处处连续可微的，也就是说，出现在下列微分等式中的系数都是连续的：

$\mathrm{d}x = x/r\,\mathrm{d}r + x/s\,\mathrm{d}s\;\mathrm{d}y = y/r\,\mathrm{d}r + y/s\,\mathrm{d}s,$

既然u关于(x, y)连续可微，那么根据微分教计算规则，得到：

$\partial u/\partial r = \partial u/\partial x * \partial x/\partial r + \partial u/\partial y * \partial y/\partial r,$

$\partial u/\partial s = \partial u/\partial x * \partial x/\partial s + \partial u/\partial y * \partial y/\partial s$。

这些偏导数都是关于新变量(r, s)连续可微的函数。于是u也关于(r, s)连续可微，因而得到：

$\mathrm{d}u = \partial u/\partial r\,\mathrm{d}r + \partial u/\partial s\,\mathrm{d}s = \partial u/\partial x\,\mathrm{d}x + \partial u/\partial y\,\mathrm{d}y$。

这表明微分形式对于x, y为任何连续可微的函数都成立。这称为（一阶）微分的形式不变性。

变量替换规定了一个坐标平面上的动点$P(x, y)$随着另一坐标平面上的动点$Q(r, s)$而变动，因而定义了一个函数$T : P = T(Q)$。这样得到一个矩阵方程：

【整理者注：不知道作者打算把雅可比矩阵写成什么格式，就不添加了。】

这里，偏导数所形成的矩阵称为雅可比矩阵。它是微分向量的系数矩阵，相当于一元函数情形的微分系数或导数。

如果动点P是在一个三维坐标空间(r, s, t)中，则函数应是三元的：$x = \varphi(r, st), y = \psi(r, s, t)$,雅可比矩阵则是：

【整理者注：不知道作者打算把雅可比矩阵写成什么格式，就不添加了。】

以此类推，一元函数微分的主要定理都能推广到多元微分中。

7.5 重积分

一元函数的定积分，作为黎曼积分和的极限，推广到二元函数几乎是直接的。只不过把积分区间换成了两个区间$X(\alpha \leq x \leq A)$和$Y(b \leq y \leq B)$，它们的乘积$R = X \times Y$是包含有界闭区域S的（各边平行于坐标轴的）最小的矩形。对于R上不属于S的点，取函数值为0,并仿照一元的情形作黎曼和数：

$$S_\Delta = \sum f(\zeta_i, \eta_j)\Delta x_i \Delta y_j, (\zeta_i, \eta_j) \in \Delta x_i * \Delta y_j (i, j = 1 \dots n, m)$$

分划(Δ)的细密程度由全部$\Delta x_i, \Delta y_j$的最大值$\|\Delta\|$来度量。于是，可以像一元的情形一样来定义二重积分：

$$\int\int_S f(x, y)\,\mathrm{d}s = \int\int_R f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \lim S_\Delta (\|\Delta\| \to 0)$$

如果这个极限存在，就说函数f 在区域S上是可积的。

可积的一个充分必要条件仍然是:函数有界并且几乎处处连续（即不连续点形成一个零测度集合）。不过，这里的零测度集合，作为平面上的点集，是指能用总面积任意小的矩形序列覆盖住。

在可积的前提下，二重积分可以写成：

$$\int\int_S f(x, y)\,\mathrm{d}s = \int_b^B \int_a^A f(x, y)\,\mathrm{d}x$$

内层积分以y为参变量，在不可积（因而相应的y值形成一个一维零测度集合）时算作0。

面积微分 $\mathrm{d}R = \mathrm{d}x\,\mathrm{d}y$,作为一个微小矩形的面积，在坐标变换之下成为一个以向量$\{\partial x/\partial r\,\mathrm{d}r, \partial y/\partial r\,\mathrm{d}r\}$和$\{\partial x/\partial s\,\mathrm{d}s, \partial y/\partial s\,\mathrm{d}s\}$ 为一对邻边的平行四边形的面积。

所以有二重积分的换元公式：

$$\int\int_S f(x, y)\,\mathrm{d}s = \int\int_{r*s} f(x, y, z)(EG - F^2)^{1/2}\,\mathrm{d}r\,\mathrm{d}s$$

7.6 三维空间的曲面积分

二重积分可以推广到三维空间中的一块曲面S上,只要这曲面是光滑的，即其上的动点$P(x, y, z)$的坐标能够表示成某一平面矩形$S = r * s(\alpha \leq r \leq A, b \leq s \leq B)$上的连续可微的函数，而以(r,s)作为P 的一种新的坐标（曲面坐标）。这里S的微小矩形$(\Delta r) \times (\Delta s)$对应着$S$上的微小曲面四边形$\Delta S$，后者的面积关于前者的面积。$\Delta r \Delta s$ 的线性主要部分便是曲面的面积微分 $\mathrm{d}S$。它等于以切线向量$\{\partial x/\partial r\,\mathrm{d}r, \partial y/\partial r\,\mathrm{d}r, \partial z/\partial r\,\mathrm{d}r\}$和$\{\partial x/\partial s\,\mathrm{d}s, \partial y/\partial s\,\mathrm{d}s, \partial z/\partial s\,\mathrm{d}s\}$ 为一对邻边的平行四边形的面积：$\mathrm{d}s = (EG - F^2)^{1/2}$，其中：

$$E = (\partial x/\partial r)^2 + (\partial y/\partial r)^2 + (\partial z/\partial r)^2$$

$$F = \partial x/\partial r * \partial x/\partial s + \partial y/\partial r * \partial y/\partial s + \partial z/\partial r * \partial z/\partial s$$

$$G = (\partial x/\partial s)^2 + (\partial y/\partial s)^2 + (\partial z/\partial s)^2$$

从而面积分能够表示成二重积分：

$$\int\int_{s} f(x,y,z)\,\mathrm{d}s = \int\int_{r*s} f(x,y,z)(EG-F^2)^{1/2}\,\mathrm{d}r\,\mathrm{d}s$$

曲面S可以是逐片光滑的,积分便取为各片上的积分之和。

如果是三维空间的曲线积分，类似地考虑空间中一条光滑的（或逐段光滑的）曲线C上关于弧长的微分 $\mathrm{d}s$ 的积分$\int_c f(x,y,z)\,\mathrm{d}s$，则有

$$\int_c f(x,y,z)\,\mathrm{d}s = \int_a^b f(x,y,z)\sqrt{(\mathrm{d}x/\mathrm{d}t)^2+(\mathrm{d}y/\mathrm{d}t)^2+(dz/\mathrm{d}t)^2}$$

这就与一个直线段$a \le s \le b$上的定积分没区别了。

实际上多元定积分在概念上的各种推广，在计算上仍都能回到定积分。

7.7　牛顿-莱布尼茨公式推广

我们知道，一元微积分之所以成立，就是靠牛顿-莱布尼茨公式。

多元微积分想成立，也得有这种把微分和积分联系起来的公式。

在一元微积分中，根据牛顿-莱布尼茨公式，定积分是微分之逆，在多元微积分中，这个定理仍然是成立的。

二重积分推广：设函数$f(x,y)$在矩形区域$D = |(x,y)|(a \le x \le b, c \le y \le d)$上连续，如果存在一个二元函数$F(x,y)$,使得$\partial^2 F(x,y)/\partial x\partial y = f(x,y)$，则二重积分$\int\int_D f(x,y)\,\mathrm{d}x\,\mathrm{d}y = F(b,d)-F(b,c)$ 更多重积分也有类似公式。

对曲线积分，也有类似公式。设D为单连通区域，$P(x,y)$和$Q(x,y)$在区域D上有连续的一阶偏导数，若存在一个二元函数$u(x,y)$，使得

$$\mathrm{d}u(x,y) = P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y$$

在区域D中任意取两个点A,B，则对连接A,B的任意一条光滑曲线L，都有:

$$\int_L P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y = u(B)-u(A)$$

另外，必须熟悉的还有斯托克斯公式（格林公式），奥斯特罗格拉茨基公式（高斯公式）等等，只是这些公式没法在豆瓣显示，有兴趣的自己去查书。

多元积分的计算技巧主要是变量替换，教科书中有大量人类积累下来的变量替换的技巧例子，可以通过多做习题，积累下自己的计算技巧，熟能生巧，培养出自己强大的计算能力。

显然，介绍的都是最古典微积分在多元上函数上的推广，现代教科书没有这么复杂，简单明了，例如定义多元函数可微，一般是:

设f是从欧几里得空间Ω（或者任意一个内积空间）中的一个开集射到R^m 的一个函数。对于Ω中的一点x及其在Ω中的邻域Λ中的点x+h。如果存在线性映射A使得对任意这样的$x+h$，$\lim|f(x+h)-f(x)-A(h)|/|h| = 0(h \to 0)$,那么称函数$f$ 在点x处可微。线性映射A叫做f 在点x处的微分，记作 $\mathrm{d}f_x$。

如果f 在点x处可微，那么它在该点处一定连续，而且在该点的微分只有一个。

当函数在某个区域的每一点x都有微分 $\mathrm{d}f_x$时，可以考虑将x映射到 $\mathrm{d}f_x$的函数：$\mathrm{d}f : x \to \mathrm{d}f_x$，这个函数一般称为微分函数。

而且利用一元微分性质，可以证明：如果f是线性映射，那么它在任意一点的微分都等于自身。

在R^n（或定义了一组标准基的内积空间）里，函数的全微分和偏导数间的关系可以通过雅可比矩阵刻画。

也可以证明如下结论：

可微的必要条件：如果函数f 在一点x_0处可微，那么雅克比矩阵的每一个元素都存在，但反之不真。

可微的充分条件：如果函数f 在一点x_0 的雅克比矩阵的每一个元素都在x_0连续，那么函数在这点处可微，但反之不真。

8 级数

级数主要两个用途，一个是构造新函数，一个是表示、逼近已知函数（主要用于函数的近似计算）。

在微积分中，会涉及一些初等函数之外的函数，一般都是用级数表达的，因为他们的级数形式，便于了解它们的性质。

级数的基本工具是泰勒级数（用有限项的多项式近似表示函数）和三角级数（傅利叶级数，表达周期性函数），级数主要用于连续函数的局部逼近和整体逼近，当然从逻辑上来讲，可以用无穷多项的多项式来准确地表示一个函数，这就是幂级数。利用函数的幂级数展开式，对研究函数的性质和计算都有着非常重要的作用。

当然，能表示成幂级数的函数必须具备任意阶可微的条件，这对于有些性质较差的函数（如分段函数），我们就不能展开成幂级数，此时付立叶级数却能满足这样的函数的展开。

级数理论的基础是极限，级数是一个无限求和的过程，它与有限求和有着根本的不同，即参与了极限运算，把极限及其运算性质移植到级数中去，就形成了级数的一些独特性质。

所以级数的第一个重要概念是收敛性（也即存在极限）。此外，级数的运算、函数项级数的一致收敛性、一致收敛级数的分析性质、函数的幂级数展开、函数的付立叶级数展开都是级数理论的基本内容。

8.1 数列级数

将数列u_n 的项$u_1, u_2, \ldots, u_n$ 用加号连接起来的函数就称数项级数简写为$\sum u_n$ ，记$S_n = \sum u_n$,如果当$n \to \infty$ 时，Sn这个数列有极限，则说级数收敛，并以S为其和,否则就说级数发散。

级数收敛的柯西准则：$\sum u_n$ 收敛$\Leftrightarrow$任意给定正数ε ，必有自然数N，当$n > N$，对一切自然数p，有$|u[n+1] + u[n+2] + , + u[n+p]| < \varepsilon$。即级数充分靠后的任意一段和的绝对值可任意小。

如果每一$u_n \geq 0$（或$u_n \leq 0$），则称$\sum u_n$ 为正（或负）项级数，正项级数与负项级数统称为同号级数。正项级数收敛的充要条件是其部分和序列S_m 有上界。

有无穷多项为正，无穷多项为负的级数称为变号级数，其中最简单的是形如$\sum[(-1)^{n-1}] * u_n(u_n > 0)$的级数，判别这类级数收敛的基本方法是莱布尼兹判别法：若$u_n \geq u_{n+1}$，对每一$n \in N$ 成立，且当$n \to \infty$ 时$\lim u_n = 0$，则交错级数收敛。

对于一般的变号级数如果有$\sum |u_n|$收敛，则称变号级数绝对收敛。如果只有$\sum u_n$ 收敛，但是$\sum |u_n|$发散，则称变号级数条件收敛。

（例如$\sum[(-1)^{n-1}] * (1/n^2)$绝对收敛，而$\sum[(-1)^{n-1}] * (1/n)$只是条件收敛）。

对条件收敛的级数有一个重要性质，也即黎曼定理：一个条件收敛的级数，在其项经过适当的排列之后，可以收敛到一个事先任意指定的数；也可以发散到$+\infty$ 或$-\infty$；也可以没有任何的和。

8.2 函数级数

如果级数的每一项依赖于变量x，x在某区间I内变化,即$u_n = u_n(x), x \in I$，$\sum u_n(x)$称为函数级数。

若$x = x_0$ 使数项级数$\sum u_n(x_0)$收敛，就称x_0 为收敛点，由收敛点组成的集合称为收敛域，若对每一$x \in I$，级数$\sum u_n(x)$都收敛，就称I为收敛区间。

显然，函数级数在其收敛域内定义了一个函数，称之为和函数$S(x)$，即$S(x) = \sum u_n(x)$如果满足更强的条件，$S_m(x)$在收敛域内一致收敛于$S(x)$。

$\sum a_n(x - x_0)^n$叫幂级数，收敛域是一个以x_0 为中心的区间（不一定包括端点），并且在一定范围内具有类似多项式的性质，在收敛区间内能进行逐项微分和逐项积分等运算。例如幂级数$\sum(2x)^n/x$ 的收敛区间是[-1/2,1/2]，幂级数$\sum(x-2)^n/n^2$的收敛区间是[1,3]，而幂级数$\sum x^n/n!$在实数轴上收敛。

不过实际上常用的级数是傅里叶级数（三角函数构成的级数），傅利叶级数的收敛范围一般很复杂，研究它需要对实变函数论、调和分析和泛函分析知识。所以真的理解并掌握傅利叶变换，不熟悉实变函数是没法入门的。

函数级数一致收敛定义：在一个集合C上一致地收敛到它的和函数$s(x)$，是指对任意$\varepsilon > 0$，对于每一个正数级数都存在一个自然数N（不依赖于x），使得当$m > N$ 时$|s(x) - s_m(x)| = |r_m(x)| < \varepsilon$，对于一切属于$C$ 的x都成立。

这时级数的和函数$s(x)$是一个无限项的和，便可在整个集合C上通过特征性质继承有限项和的一些分析性质：

★逐项积分定理：设函数级数级数在有限闭区间$\alpha \leq x \leq b$上一致地收敛，若级数的各项$s_N(x)$都连续，则级数的和也连续并且可以逐项积分。

★逐项微分定理：通过微分与积分的互逆关系（微积分基本定理）能够把上述定理转变成逐项微分的形式：设函数级数级数在区间$\alpha < x < b$内收敛，各项都具有连续的导数，若逐项取导数所得的级数在该区间内一致收敛，则原级数的和也具有连续的导数并且可以逐项微分。

8.3 函数级数收敛判定

显然下面一个主要问题是函数级数的收敛问题。因为一个函数级数在其收敛范围内代表一个函数，即它的和$\sum u_n(x)(n = 1, \ldots \infty) = u(x)$，当和是有限项时（$\sum u_n(x)(n = 1, \ldots M)$）,这个级数和就是这个$u(x)$ 函数逐步逼近定义的一种方式。

在函数级数收敛研究过程中，经过约200年，才发现一致收敛概念的价值：这种级数展开在收敛区间内可以逐项微分和积分并且收敛。

级数在逐项取绝对值之后就成为正项级数，显然可以依一致收敛性进行比较，特别是用一个常数级数进行比较，便有M判别法。

M判别法（魏尔斯特拉斯判别法）：假设$\{u_n\}$是定义在集合C内的一个实数或复数函数的数列，并存在正的常数M_n，使得$|u_n(x)| \leq M_n$

对于所有的$n \geq 1$和C内所有的x成立。进一步假设级数$\sum M_n (n = 1, \ldots \infty)$收敛。那么级数$\sum u_n(x) (n = 1, \ldots \infty)$在$C$内一致收敛。（可由常数项级数收敛的柯西准则证明）。

8.4 泰勒级数

我们常用的级数函数之一是泰勒级数。

泰勒级数定义：如果$f(x)$在点$x = x_0$ 具有任意阶导数，则幂级数：

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n = f(x_0)+f'(x_0)*(x-x_0)+\frac{f^{(2)}(x_0)}{2!}(x-x_0)^2+\ldots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

称为$f(x)$ 在点x_0 处的泰勒级数。

在上述定义中，取$x_0 = 0$，得到的级数$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}(x)^n$称为麦克劳林级数。函数$f(x)$的麦克劳林级数是$x$的幂级数，那么这种展开是唯一的，且必然与$f(x)$的麦克劳林级数一致。

如果$f(x)$的麦克劳林级数在点的某一邻域内收敛，它不一定收敛于$f(x)$。因此，如果$f(x)$在某处有各阶导数，则$f(x)$的麦克劳林级数虽然能算出来，但这个级数能否在某个区域内收敛，以及是否收敛于$f(x)$ 还需要进一步验证。

一些函数无法被展开为泰勒级数，因为那里存在一些奇点。

定理一：设函数$f(x)$在x_0 的某个邻域$N(x_0, \delta_0)$内具有任意阶导数，则函数$f(x)$在该邻域内能展开成泰勒级数的充要条件是的泰勒公式中的余项$R_n(x)$满足$\lim R_n(x) = 0, x \in N(x_0, \delta_0)$

定理二：如果$f(x)$在区间$(-R + x_0, R + x_0)$能展开成泰勒级数

$$\sum \frac{f^{(n)}(a)}{n!}(x-a)^n \ (n = 0, \ldots \infty)$$

则右端的幂级数是惟一的。

泰勒级数的重要性质是在研究幂级数收敛过程中得到的：可以严密证明幂级数在其收敛区间内展开式是唯一的，也即幂级数能够完全代表它的和函数参加分析运算（同时也证明了三角级数展开式不具有唯一性，所以三角函数的收敛集非常样复杂，这就是后来研究三角级数收敛性的学科调和分析能够成为数学主要学科的理由：问题复杂）。

由于幂级数可以逐项微分任意多次，所以幂级数本身就是它的和函数在收敛区间中心处的泰勒级数。所以一个泰勒级数的系数不一定要单纯通过累次微分级数而可以通过某些幂级数的分析运算来求得（因为微分次数越多计算越复杂）。

由于幂级数的求导和积分可以逐项进行，因此求和函数相对比较容易。这是泰勒级数最大的用处：简化计算。

同时，在复变函数中，一个解析函数可被延伸为一个定义在复平面上的一个开区域上的泰勒级数，这样可以简化和拓展解析函数定义方式。

不过在工程中，泰勒级数主要用来近似计算函数的值。

必须强调一点是，对于一些无穷可微函数$f(x)$，虽然它们的展开式收敛，但是并不等于$f(x)$。例如，分段函数：$f(x) = e^{-1}/x^2$，当$x \neq 0$ 且$f(0) = 0$，则当$x = 0$所有的导数都为零，则这个$f(x)$在$x = 0$处的泰勒级数为零，且其收敛半径为无穷大，虽然这个函数f 仅在$x = 0$ 处为零。

下面给出几个常见函数在$x = 0$处的泰勒级数，即麦克劳林级数。

指数函数：$e^x = \sum_{n=0}^{\infty} x^n/n!$

自然对数：$\ln(x + 1) = \sum_{n=0}^{\infty}(-1)^n + 1/n * x^n$，$x \in (-1, 1]$

几何级数：$1/(1 - x) = \sum_{n=0}^{\infty} x^n$，$|x| < 1$

正弦函数：$\sin x = \sum_{n=0}^{\infty}(-1)^n/(2n + 1)! * x^{2n+1}$

余弦函数：$\cos x = \sum_{n=0}^{\infty}(-1)^n/(2n)! * x^{2n}$

8.5 傅利叶级数

傅利叶级数或者傅利叶变换，是工程师的大杀器，对搞信号分析，模式识别的工程师，基本上就是居家旅游，吃饭睡觉的唯一工具。

由于傅利叶级数涉及很多波形图，豆瓣不支持，只能直观描述，有兴趣的去查教科书，可能才能清楚我说的是什么。

傅立叶贡献：猜想周期函数都可以展开为常数与一组具有共同周期的正弦函数和余弦函数之和。（但是未能严格证明，拉格朗日就反对他的论文发表，认为不能三角级数表达梯形或箱型周期函数。后来狄利赫里证明了三角级数在一定条件下的收敛唯一性，并用级数连续逼近可以表达梯形或箱型周期函数）。

傅利叶级数展开式中，常数表达的部分称为直流分量，最小正周期等于原函数的周期的部分称为基波或一次谐波，最小正周期的若干倍等于原函数的周期的部分称为高次谐波。因此高次谐波的频率必然也等于基波的频率的若干倍，基波频率N倍的波称为N次谐波，是$N - 1$次泛音。不管几次谐波，他们都是正弦波。正弦波是基本波形。

所以简单说：傅利叶级数就是周期函数展开为一个三角级数，例如：

给定一个周期为T的函数$f(t) = a_0 + \sum_{n=1}^{\infty}[a_n \cos(n\omega t) + b_n \sin(n\omega t)]$，$\omega_0 = 2\pi/T$，$\omega t = 2\pi t/T$

根据欧拉公式，也可以等价表示为：$f(t) = \sum_{k=-\infty}^{\infty} a_k * e^{jk(2\pi/T)t}$ j为虚数单位,其中$a_k = 1/T * \int_T f(t) * e^{-jk(2\pi/T)t}\,\mathrm{d}t$,

其中$e^{jk(2\pi/T)t}$是周期为T的函数，所以k 取不同值时的周期信号具有谐波关系（即它们都具有一个共同周期T）。$k = 0$时，对应的这一项称为直流分量，$k = 1$时具有基波频率，称为一次谐波或基波，类似的有二次谐波，三次谐波等等。

欧拉公式：

$e^{jx} = \cos x + j \sin x$

$\cos x = [e^{jx} + e^{-jx}]/2$

$\sin x = [e^{jx} - e^{-jx}]/(2j)$

傅里叶最大的贡献是猜想了傅利叶级数的性质，而严格证明了傅利叶级数的收敛性则是狄利赫里。

狄利赫里定理：满足狄利赫里条件的周期函数表示成的傅里叶级数都收敛。狄利赫里条件如下：

在任何周期内，$f(t)$须绝对可积；

傅里叶级数在任一有限区间中，$f(t)$只能取有限个最大值或最小值；
在任何有限区间上，$f(t)$只能有有限个第一类间断点。

定理结论是：满足狄利赫里条件的周期函数都可以展开为正弦函数和余弦函数的级数和，并且这个展开是收敛到唯一周期函数的。这是傅利叶变换的基础定理。

既然傅利叶猜想周期函数能够展开成三角函数的级数，那么三角函数的性质就很重要。三角函数最最重要性质是正交性，因为这是证明傅利叶级数收敛的唯一条件。

正交性定义：两个不同向量正交是指它们的内积为0。（这也就意味着这两个向量之间没有任何相关性，例如，如果两个函数$\psi_1(r)$和$\psi_2(r)$满足条件：$\int \psi_1(r)\psi_2(r)\,\mathrm{d}\tau = 0$,则称这两个函数相互正交。在三维欧氏空间中，互相垂直的向量之间是正交的。事实上，正交是垂直在数学上的的一种抽象化和一般化）。

内积定义：设向量$A = [a_1, a_2, \dots a_n]$，$B = [b_1, b_2 ... b_n]$，则向量$A$和$B$的内积表示为：

$A \cdot B = a_1 \times b_1 + a_2 \times b_2 + ,, + a_n \times b_n$
$A \cdot B = |A| \times |B| \times \cos\theta$
$|A| = (a_1^2 + a_2^2 + ... + a_n^2)^(1/2)$；
$|B| = (b_1^2 + b_2^2 + ... + b_n^2)^(1/2)$。

其中，$|A|$ 和$|B|$ 分别是向量A和B的模，θ是向量A和向量B的夹角（$\theta \in [0, \pi]$）。

若B为单位向量，即$|B| = 1$时，$A \cdot B = |A|\cos\theta$，表示向量$A$在B方向的投影长度。向量$A$为单位向量时同理。　当且仅当向量$A$与$B$垂直时，$A \cdot B = 0$。

显然，学过线性代数都知道，一组n个互相正交的向量必然是线形无关的，所以必然可以张成一个n维空间，也就是说，空间中的任何一个向量可以用它们来线性表出。

函数的正交是向量正交正交的推广,函数可看成无穷维向量。

所谓三角函数系$\{1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx, \dots\}$在区间$[-\pi, \pi]$上正交,就是指在三角函数系中任何不同的两个函数的乘积在区间$[-\pi, \pi]$上的积分等于0,即

$\int_{-\pi}^{\pi} \cos(nx)\,\mathrm{d}x = 0$
$\int_{-\pi}^{\pi} \sin(nx)\,\mathrm{d}x = 0$
$\int_{-\pi}^{\pi} \sin(kx) * \cos(nx)\,\mathrm{d}x = 0$；任意$k$和$n$
$\int_{-\pi}^{\pi} \cos(kx(* \cos(nx)\,\mathrm{d}x = 0$；$(k, n = 1, 2, 3, \dots, k \neq n)$
$\int_{-\pi}^{\pi} \sin(kx) * \sin(nx)\,\mathrm{d}x = 0$；$(k, n = 1, 2, 3, \dots, k \neq n)$
$\int_{-\pi}^{\pi} \cos(mx) * \cos(nx)\,\mathrm{d}x = 0$；$(m \neq n)$
$\int_{-\pi}^{\pi} \sin(nx) * \sin(nx)\,\mathrm{d}x = \pi$
$\int_{-\pi}^{\pi} \cos(nx) * \cos(nx)\,\mathrm{d}x = \pi$

有了上述性质，傅利叶级数的展开就能大幅简化。

三角函数的另外一个重要性质是有奇偶性。

如果对于函数$f(x)$的定义域内任意一个x，都有$f(-x) = -f(x)$，那么函数$f(x)$就叫做奇函数；如果对于函数$f(x)$的定义域内任意的一个x，都有$f(x) = f(-x)$,那么函数$f(x)$就叫做偶函数。奇函数可以表示为正弦级数，偶函数则可以表示成余弦级数。

也即奇函数$f(x) = \sum_{-\infty}^{\infty} b_k \sin(kx)$；偶函数$g(x) = a_0/2 + \sum_{-\infty}^{\infty} a_k \cos(kx)$。

这些公式用欧拉公式，就可以很容易从上面傅里叶级数的公式中导出。

其实利用函数正交性，可以证明更一般的定理，那就是广义傅利叶技级数的收敛性。

广义傅里叶级数是对一切正交函数系定义的，类比三角函数定义的傅利叶级数。

定义：任何正交函数系$\{g(x)\}$，如果定义在$[a,b]$上的函数$f(x)$只具有有限个第一类间断点，那么如果$f(x)$满足封闭性方程：$\int_a^b f^2(x)\,\mathrm{d}x = \sum_1^\infty C_k^2$，那么级数$\sum_1^\infty C_k g_k(x)$必然收敛于$f(x)$，其中：

$$c_n = \int_a^b f(x)g_n(x)\,\mathrm{d}x$$

事实上，无论级数$\sum_1^\infty C_k g_k(x)$是否收敛，总有：$\int_a^b f^2(x)\,\mathrm{d}x \geq \sum_1^\infty C_k^2$（贝塞尔(Bessel) 不等式）。

这个性质经常用，可以大幅简化问题。

8.6 傅利叶变换和调和分析简介

傅利叶变换在物理学、电子类学科、数论、组合数学、信号处理、概率论、统计学、密码学、声学、光学、海洋学、结构动力学等领域都有着广泛的应用（例如在信号处理中，傅里叶变换的典型用途是将信号分解成频率谱——显示与频率对应的幅值大小）。

由于傅利叶变换的巨大用途（目前尚未有任何数学工具在实际工程和科学应用上可以与之相提并论），下面稍微多说几句。

★傅里叶变换定义

简单说，傅立叶变换是一种分析周期函数（例如信号）的方法，它可分析周期函数的频率成分或时变成分，也可用这些成分合成函数（或信号）。（虽然许多波形可作为函数（信号）的成分，比如正弦波、方波、锯齿波等，但是傅立叶变换用正弦波作为信号的成分，因为其容易计算）。

●连续型傅利叶变换

常用的主要是连续型傅利叶变换。

连续型傅利叶变换的定义：$f(t)$ 是t的周期函数，如果t满足狄利赫里条件：在一个以$2T$为周期内$f(x)$连续或只有有限个第一类间断点），且$f(x)$单调或可划分成有限个单调区间，则F(x)以2T为周期的傅里叶级数收敛，和函数$S(x)$也是以$2T$为周期的周期函数，且在这些间断点上，函数是有限值；在一个周期内具有有限个极值点；绝对可积。

则有$F(\omega) = \int_{-\infty} \infty f(t) * e^{-i\omega t}\,\mathrm{d}t$称为积分运算$f(t)$的傅立叶变换，即将频率域的函数表示为时间域的函数。

$f(t) = 1/2\pi * \int_{-\infty} \infty F(\omega) * e^{-i\omega t}\,\mathrm{d}\omega$叫做$F(\omega)$的傅立叶逆变换，即将时间域的函数表示为频率域的函数。

$F(\omega)$叫做$f(t)$ 的像函数，$f(t)$ 叫做$F(\omega)$的像原函数。$F(\omega)$是$f(t)$ 的像。$f(t)$是$F(\omega)$原像。

连续形式的傅里叶变换其实是傅里叶级数的推广，因为积分其实是一种极限形式的求和算子而已。对于周期函数，它的傅里叶级数表示被定义为：$f(t) =$

$\sum_{-\infty}^{\infty} F_n * e^{jn(2\pi/T)t}$，其中$T$ 为函数的周期，F_n为傅里叶展开系数，$F_n = 1/T * \int_{-T/2}^{T/2} f(t) * e^{-in(2\pi/T}\,\mathrm{d}t$，

对于实值函数，函数的傅里叶级数可以写成：$f(t) = a_0 + \sum_{n=1}^{\infty}[a_n \cos(n\omega t) + b_n \sin(n\omega t)]$，$\omega_0 = 2\pi/T$，$\omega t = 2\pi t/T$ 其中a_n和b_n是实频率分量的振幅。

当$f(t)$ 为奇函数（或偶函数）时，其余弦（或正弦）分量为零，而可以称这时的变换为余弦变换（或正弦变换）。

●离散时间傅里叶变换

针对的是定义域为Z的数列。设$\{x_n\}[-\infty,\infty]$为某一数列，则其离散时间傅里叶变换被定义为:

$$X(\omega) = \sum_{-\infty}^{\infty} x_n * e^{-i\omega n}$$

相应的逆变换为

$$x_n = 1/2\pi * \int_{-\infty}^{\infty} X(\omega) * e^{-i\omega t}\,\mathrm{d}\omega$$

离散时间傅里叶变换在时域上离散，在频域上则是周期的，它一般用来对离散时间信号进行频谱分析。

●离散傅里叶变换

离散函数且满足有限性或周期性条件，序列$\{x_n\}[n = 0,\ldots,N-1]$ 的离散傅里叶变换为

$$x[k] = \sum_{n=0}^{N-1} x_n * e^{-i2\pi kn/N}$$

其逆变换为：

$$x_n = 1/N * \sum_{n=0}^{N-1} X[k] * e^{i2\pi kn/N}$$

傅里叶变换可以将计算复杂度降低（这个变换在数字电路计算，信号处理等等行业是十分实用且重要的方法）。

★傅利叶变换基本性质

●傅里叶变换具有线性性质：假设函数$f(x)$和$g(x)$的傅里叶变换都存在，a和b为任意常系数，则有$F[af + bg] = aF[f] + bF[g]$

●尺度变换性质：若函数$f(x)$的傅里叶变换为(ω)，则对任意的非零实数a，函数$f_a(x = f(ax)$的傅里叶变换$F_a(\omega)$存在，且等于$F(a\omega) = 1/|a| * F(\omega/a)$。也即当$a > 0$时，若将$f(x)$ 的图像沿横轴方向压缩a 倍，则其傅里叶变换的图像将沿横轴方向展宽a倍，同时高度变为原来的$1/a$。对于$a < 0$时，傅里叶变换的图像关于纵轴做镜像对称。

●对偶性质：若函数$f(x)$的傅里叶变换为(ω)，则存在$F(x)$的傅利叶变换$= 2\pi f(-\omega)$

●平移性质：若函数$f(x)$的傅里叶变换为(ω)，则对任意实数a，函数$f_a(x = f(x) * e^{iax}$也存在傅里叶变换，且其傅里叶变换$F_a(\omega) = F(\omega - a)$。也即$F_a(\omega)$可由$F(\omega)$ 向右平移a 得到。

●微分关系：若函数$f(x)$的傅里叶变换为(ω)，且其导函数$f'(x)$ 的傅里叶变换存在，则有$f'(x)$的傅利叶变换$= i * \omega * F(\omega)$，也即导函数的傅里叶变换等于原函数的傅里叶变换乘以因子$i * \omega$。更一般地，若$f(x)$的n阶导数的傅里叶变换

存在，则$f(x)$的n 阶导数的傅里叶变换$= (i*\omega)^n * F(\omega)$，即$n$阶导数的傅里叶变换等于原函数的傅里叶变换乘以因子$(i*\omega)^n$。

★傅利叶变换基本定理

●时域卷积定理

若函数$f(x)$的傅里叶变换为$F(\omega)$，若函数$g(x)$的傅里叶变换为$G(\omega)$，若函数$f(x)$,$g(x)$都在R上绝对可积，则卷积函数

$$f*g(x)=\int_{-\infty}^{\infty} f(x-t)*g(t)\,\mathrm{d}t$$

的傅里叶变换存在，且$f*g(x)$的傅利叶变换$= F(\omega)*G(\omega)$

●频域卷积定理

若函数$f(x)$的傅里叶变换为$F(\omega)$，若函数$g(x)$的傅里叶变换为$G(\omega)$，则有$f(x)*g(x)$的傅利叶变换$= 1/2\pi*[F(\omega)*G(\omega)]$

★傅利叶变换用途广泛的原因

主要原因是傅利叶变换有下面这些优点。

●傅里叶变换是线性算子,若赋予适当的范数,它还是酉算子；

●傅里叶变换属于谐波分析；

●傅里叶变换的逆变换容易求出,而且形式与正变换非常类似；

●正弦基函数是微分运算的本征函数,从而使得线性微分方程的求解可以转化为常系数的代数方程的求解。在线性时不变杂的卷积运算为简单的乘积运算,从而提供了计算卷积的一种简单手段；

●离散形式的傅里叶的物理系统内（线性时不变）,频率是个不变的性质,从而系统对于复杂激励的响应可以通过组合其对不同频率正弦信号的响应来获取；

●卷积定理指出：傅里叶变换可以化复杂的卷积运算为简单的乘积运算，从而提供了计算卷积的一种简单手段；

●离散形式的傅立叶变换可以利用快速傅里叶变换算法（FFT)。

★从信号分解的几何直观角度（波形图）来简单解释一下傅利叶变换的思想

傅利叶变换核心思想是用正弦曲线来代替原来的曲线而不用方波或三角波来表示。

原因在于分解信号的方法是无穷的，但分解信号的目的是为了更加简单地处理原来的信号。用正余弦来表示原信号会更加简单，因为正余弦拥有原信号所不具有的性质：正弦曲线保真度。一个正弦曲线信号输入后，输出的仍是正弦曲线，只有幅度和相位可能发生变化，但是频率和波的形状仍是一样的。且只有正弦曲线才拥有这样的性质，正因如此我们才不用方波或三角波来表示。

为什么选择三角函数而不用其他函数进行分解？因为很多现象可以抽象成一个线性时不变系统（也即输入输出信号满足线性关系，而且系统参数不随时间变换，无论用微分方程还是传递函数或者状态空间描述都可以），而且一个正弦曲线信号输入后，输出的仍是正弦曲线，只有幅度和相位可能发生变化，但是频率和波的形状仍是一样的。也就是说正弦信号是系统的特征向量，同时指数信号也是系统的特征向量（表示能量的衰减或积聚，衰减或者扩散现象大多是指数形式的，或者既有波动又有指数衰减，也即e^{a+ib}形式），所以除了指数信号和正弦信号以外的其他波形都不是线性系统的特征信号。

由于正弦信号是很多线性时不变系统的特征向量，于是傅里叶变换就有了用武之地。对于更一般的线性时不变系统，复指数信号(表示耗散或衰减)是系统的特征向量，于是拉普拉斯变换就有了用武之地。

显然，傅里叶级数和傅里叶变换就能处理特征值与特征向量的问题，这样用正余弦来表示原信号会更加简单，因为正余弦拥有原信号所不具有的性质：正弦曲线保真度。且只有正弦曲线才拥有这样的性质。

这也解释了傅利叶变换的强大用途的原因：因为正弦量(或复指数)是特征向量。

★傅里叶变换的推广

从数学的角度理解积分变换就是通过积分运算，把一个函数变成另一个函数。也可以理解成是算内积，然后就变成一个函数向另一个函数的投影：$F(s)=\int_a^b f(t)*K(s,t)\,\mathrm{d}t$，$K(s,t)$是积分变换的核(Kernel)。

当选取不同的积分域和变换核时，就得到不同名称的积分变换（也即向核空间投影，将原问题转化到核空间。所谓核空间，就是这个空间里面装的是核函数）。

当然，选取什么样的核主要看面对的问题有什么特征。不同问题的特征不同，就会对应特定的核函数。把核函数作为基函数。将现在的坐标投影到核空间里面去，问题就会得到简化。之所以叫核，是因为这是最核心的地方。至于常用傅里叶变换和拉普拉斯变换是因为复指数信号才是描述这个现实世界的特征函数。

★傅利叶变换使用的一个例子

图像的频率是表征图像中灰度变化剧烈程度的指标，是灰度在平面空间上的梯度。如：大面积的沙漠在图像中是一片灰度变化缓慢的区域，对应的频率值很低；而对于地表属性变换剧烈的边缘区域在图像中是一片灰度变化剧烈的区域，对应的频率值较高。设f是一个能量有限的模拟信号，则其傅里叶变换就表示f的谱。

从物理效果看，傅里叶变换是将图像从空间域转换到频率域，其逆变换是将图像从频率域转换到空间域。换句话说，傅里叶变换的物理意义是将图像的灰度分布函数变换为图像的频率分布函数，傅里叶逆变换是将图像的频率分布函数变换为灰度分布函数。

傅里叶变换以前，图像（未压缩的位图）是由对在连续空间（现实空间）上的采样得到一系列点的集合，用一个二维矩阵表示空间上各点，则图像可由$z=f(x,y)$来表示。由于空间是三维的，图像是二维的，因此空间中物体在另一个维度上的关系就由梯度来表示。

对图像进行二维傅里叶变换得到频谱图，就是图像梯度的分布图，当然频谱图上的各点与图像上各点并不存在一一对应的关系，即使在不移频的情况下也是没有。

从傅里叶频谱图上看到的明暗不一的亮点，实际上图像上某一点与邻域点差异的强弱，即梯度的大小，也即该点的频率的大小（图像中的低频部分指低梯度的点，高频部分相反）。梯度大则该点的亮度强，否则该点亮度弱。

傅里叶变换后的频谱图，也叫功率图。

可以看出，图像的能量分布，如果频谱图中暗的点数更多，那么实际图像是比较柔和的（因为各点与邻域差异都不大，梯度相对较小），反之，如果频谱图中亮的点数多，那么实际图像一定是尖锐的，边界分明且边界两边像素差异较大的。

对频谱移频到原点以后，可以看出图像的频率分布是以原点为圆心，对称分布的。将频谱移频到圆心可以分离出有周期性规律的干扰信号（带有正弦干扰，移频到原点的频谱图上可以看出除了中心以外还存在以某一点为中心，对

称分布的亮点集合，这个集合就是干扰噪音产生的），这时可以很直观的通过在该位置放置带阻滤波器消除干扰。

图像经过二维傅里叶变换后，其变换系数矩阵表明若变换矩阵Fn原点设在中心，其频谱能量集中分布在变换系数短阵的中心附近，若所用的二维傅里叶变换矩阵Fn的原点设在左上角，那么图像信号能量将集中在系数矩阵的四个角上。这是由二维傅里叶变换本身性质决定的。同时也表明一股图像能量集中低频区域。

变换之后的图像在原点平移之前四角是低频，最亮，平移之后中间部分是低频，最亮，亮度大说明低频的能量大（幅角比较大）。

这样通过傅利叶变换，就能简化计算，识别图像特征。

★调和分析简介

傅里叶分析从诞生之日起，就围绕着“傅里叶级数究竟是否收敛于自身”这样一个中心问题进行研究，这也是调和分析的中心问题。

可以说调和分析就是傅里叶分析，调和分析是研究作为基本波形的叠加的函数或者信号的表示的数学分支。它研究并推广傅立叶级数和傅立叶变换的概念。基本波形称为调和函数，和分析因此得名。主要用途是信号处理、量子力学、模式识别、人工智能、神经科学等等。

当初傅里叶只是提出周期函数可用三角级数表示的猜想，并未证明。是狄利克雷给出了周期函数的傅里叶级数收敛于它自身的充分条件：一个周期上分段单调的周期函数的傅里叶级数，在它的连续点上必收敛于$f(x)$；如果在x点不连续，则级数的和是$(f(x+0)+f(x-0))/2$。

狄利克雷的定理表明：函数在一个周期内的分段单调性，可能导致该函数在不同区间上的不同解析表示，这自然应当把它们看做同一个函数的不同组成部分，而不是像当时人们所理解的那样，认为一个解析表达式就是一个函数。这是对函数概念的一大突破和重大贡献。

黎曼对傅里叶级数的研究也作出了贡献。黎曼在1854年《用三角级数来表示函数》论文中，引进了现在称为黎曼积分的概念及其性质，证明了如果周期函数(x)在$[0,2\pi]$上有界且可积，则当n趋于无穷时的傅里叶系数趋于0；有界可积函数的傅里叶级数在一点处的收敛性，仅仅依赖于(x)在该点近旁的性质。这是一个本质定理，现在称之为局部性原理。

海涅在1870年证明：有界函数$f(x)$可以唯一地表示为三角级数，但是由于证明不完备（因为傅里叶级数未必一致收敛，从而无法确保逐项积分的合理性，逻辑上就可能存在不一致收敛的三角级数。但这个级数确实表示一个函数），导致了康托研究函数用三角级数表示是否唯一的问题的由来：为此康托引进了点集的极限点以及导集等概念，这导致了实变函数的诞生。

魏尔斯特拉斯在1861年首次利用三角级数构造了处处不可求导的连续函数。他的这一发现震动了当时的数学界，因为长期的直观感觉使人们误认为，连续函数只有在少数一些点上才不可求导。

这个发现，直接导致了勒贝格积分和点集测度理论诞生。（勒贝格积分与勒贝格测度，现在已成为数学各分支中不可缺少的重要概念和工具）。勒贝格利用勒贝格积分和点集测度把黎曼的工作又推进了一步，得到如下重要结果：任何勒贝格可积函数的傅里叶级数，不论收敛与否，都可以逐项积分；对于$[0,2\pi]$上勒贝格平方可积的函数，帕舍伐尔等式成立（$\|x\|^2=|(x,ek)|^2$）。

连续函数的傅里叶级数，是否必处处收敛？1876年杜布瓦- 雷蒙发现，存在连续函数，它的傅里叶级数在某些点上发散；后又证明，连续函数的傅里叶级数可以在一个无穷点集上处处发散。这反面结果的发现提醒人们对傅里叶级数

的收敛性应持审慎态度。这些重要发现，导致了对傅利叶级数收敛性质的进一步探讨，结果成果越来越多，最后形成现代数学一个主要分支：调和分析。

另外一个研究傅里叶级数收敛的方向是复变函数论方法。因为傅里叶级数的指数函数表达式可以看成单位圆内的解析函数（可积函数的傅里叶级数它是复变量z的幂级数的实部。所以复变函数论是研究傅里叶级数的一个重要工具。

利用复变函数，哈代和里斯兄弟建立单位圆上H空间的理论。

50年代以前，傅里叶分析的研究领域基本上限于一维的具体空间，50年代以后的研究，逐渐向多维和抽象空间推广，例如傅立叶级数在希尔伯特空间的意义研究，以此建立了与泛函分析的一个联系；再例如基于拓扑群上的函数或测度以及由它们构成的空间或代数来研究傅利叶变换；以及考尔德伦- 赞格蒙奇异积分理论，伯克霍尔德的一般H空间理论，以及群上的傅里叶分析。

从群论的观点看，无论是周期函数还是非周期函数，它们的定义域都是拓扑群G（要在群上运用傅里叶分析方法，先就要能在群上定义傅里叶变换，外与彼得合作对一般紧李群建立了外尔-彼得定理,奠定了紧群上调和分析的基础，哈尔对满足某些条件的局部紧群证明了特殊测度（哈尔测度）的存在性），就是说，G有一个代数运算，称为群运算，以及与之相协调的极限运算，称为G的拓扑。傅里叶级数或傅里叶积分的任务，正是研究G上定义的函数$f(x)$分解为群上许多“特殊”函数（例如e或e）之和的可能性，以及通过傅里叶系数或傅里叶变换来研究自身的性质。例如，以乘法为群运算的全体正实数构成一拓扑群R，它的拓扑就是欧氏空间的拓扑，那么测度$d\mu = x\,\mathrm{d}x$就是R上的哈尔测度。可以证明对于群R上的可积函数$f(x)$的傅里叶变换收敛性。群代数、测度代数、傅里叶代数、傅里叶- 斯蒂尔杰斯代数这些是群上调和分析最主要的研究对象。

群论观点的引入，使得隐藏在周期性函数现象背后的内在联系，被揭示得更清楚更深刻了，使得调和分析内部各分支之间以及调和分析与其他学科例如泛函分析、代数学、群表示论、模形式等的联系变得更为密切。因此，群上调和分析可以说是一门既具应用价值（正如它对概率论、数论与微分方程等所起的作用所说明的）又具理论意义的综合性学科。

有兴趣的人可以去查调和分析教科书。

9 现代微积分

上面介绍的微积分是最古典的微积分，也即是18世纪19世纪的微积分，也是现在绝大多数理工科院校教授的微积分，他们特点是涉及的微分，积分和级数讨论的函数的自变量定义的区域基本都是一维的直线，二维的曲线和三维的曲面。但是现在好的大学数学系的本科学生学的微积分是现代微积分，例如哈佛。普林斯顿、耶鲁、斯坦福、MIT、加州理工等等的数学系，当然中国科大数学系本科的微积分也是现代微积分。

现代微积分的标志之一就是从流形上的微积分开始。古典微积分讨论对象–函数的定义一般都是实数数域上的一个映射。而现代微分讨论的对象是流形上的映射。

流形（Manifold）的概念最早是在1854年由Riemann 提出，简单说就是局部具有欧氏空间性质的空间，一般可以直观理解流形是把许多平直的片折弯并粘连而成。它是数域概念的推广，例如可以象数域一样定义距离（一般用测度表

示，包括面积体积等等概念），定义方向（包括向量场），定义运算（包括各种算子，变换，内积等等），以及定义流形上的微分和积分。

流形是一个几何概念，流形是任意维度的抽象空间，简单说流形包括各种维数的曲线曲面。（线段是一维的，曲面是二维的，三维空间中的所有旋转是三维的)。微积分研究其可微性。流形概念来自于物理，例如经典力学的相空间和构造广义相对论的时空模型的四维伪黎曼流形都是流形的实例。

欧几里得空间就是流形最简单的实例，像地球表面这样的球面也是一个流形。

欧几里德空间是一个特别的度量空间，定义如下：设V是实数域R上的线性空间（向量空间），若g是V上的二元实值函数，满足如下关系：

★$g(x, y) = g(y, x)$;

★$g(x + y, z) = g(x, z) + g(y, z)$;

★$g(kx, y) = kg(x, y)$;

★$g(x, x) \geq 0$，而且$g(x, x) = 0$当且仅当$x = 0$时成立。

这里x, y, z是V中任意向量，k是任意实数。

内积空间是对欧氏空间的一般化（一个线性空间定义了内积运算之后就是欧几里德空间，向量空间又称线性空间，是线性代数的中心内容和基本概念之一）。内积空间和度量空间都是泛函分析的基本研究对象。

举几个经典欧几里德空间例子：

E^n：在n维实向量空间R^n中定义内积$(x, y) = x_1y_1 + ... + x_ny_n$，则$R^n$为欧几里德空间。（任意一个$n$维欧几里德空间V等距同构于$E^n$）

设V是[0,1]区间上连续实函数全体，则V是R上线性空间，对于如下内积是欧几里德空间：(f, g)定义为fg在[0,1]区间上的积分值。

简单来讲，流形上的微积分课程首先得介绍欧几里得空间性质，包括范数、线性变换和连续映射，然后介绍可微映射及其导数和逆映射定理，然后要介绍欧几里得空间上的可积函数的特征，然后介绍微分流形特征，及其微分形式和外微分，流形上的积分，斯托克斯公式等等。由于具体内容有太多数学符号，在豆瓣无法表达，例如流行上的微分就涉及流形映射的雅可比矩阵，所以只能简单提一提，具体内容，有兴趣的可以找书来看，现在物理学和一些应用科学，例如模式识别，不懂流形微积分，基本没法玩。

就我的经历来看，在大学本科讲流形上的微积分，对大多数学生来讲毫无难度。淘汰数学系学生的两道门槛，一道是$\varepsilon - \delta$语言体系（也包括充分必要条件，逻辑完备性，反例等等概念）的充分理解和把握，一道是代数结构抽象语言（也包括同构，同态，同胚等等概念）的充分理解和把握。一般抽象能力和逻辑能力跟不上的学生，都会在这两道门槛前被淘汰。

显然根据前面介绍，我们知道微积分学的基础概念其实是无限，传统微积分把无限当成一个过程（也即维斯特拉斯用$\varepsilon - \delta$语言公理体系定义的极限的概念），但是在不断发现不连续函数，不可积函数后（其实微分方程很多解的函数都不是初等函数，很多都只能用级数表达或间断函数表达，这里面就有许多不可积或不可微，甚至连续不可微的函数，例如魏尔斯特拉斯函数就是一类处处连续而处处不可导的实值函数。魏尔斯特拉斯函数是一种无法用笔画出任何一部分的函数，因为每一点的导数都不存在，画的人无法知道每一点该朝哪个方向画。魏尔斯特拉斯函数的每一点的斜率也是不存在的。在魏尔斯特拉斯这个反例出现之前，数学家们认为除了少数一些特殊的点以外，连续的函数曲线在每一点上总会有斜率），所以传统微积分逻辑基础需要的公理体系就存在了漏洞，因为这个公理体系假设了无限是一个具体的东西，一种真实的存在，但

是很多反例在质疑这种假设。为了弥补这个漏洞，就自然出现了进一步研究实数无限性质的必要，这就是下面我们要介绍的实变函数的内容。

10 证明解存在性的逻辑价值的一个例子

对工程师来讲，能够用来计算的数学工具才是有用的，由于微积分强大的计算能力，就成为了必备工具（另外工程师常用的还有线性代数和数理统计），但是从逻辑角度来讲，证明解的存在性重要性比计算解要有价值得多，因为去计算不存在的解是无用功。举例来讲，微观经济学里面的阿罗–德布鲁边际均衡模型证明的结论：一定假设下，供需曲线一定相交于一点（也即均衡价格是存在的），这个结论是整个微观经济学的逻辑基础，没有这个存在性证明，微观经济学其实不是科学，只是假设和幻想。这个基础定理不但可以引申证明微观经济学的一些核心概念的正确性，例如边际效益递减（也即增长是有极限的），最优增长路径存在（也即著名的萨缪尔森大道定理）和经济体系（不管开放或封闭），只要资源约束条件是凸集，就有多目标非劣解等等。这些定理就是存在性定理的杰作。

当然我们非经济经济专业人士一般知道阿罗不是因为这个均衡定理，而是另外一个存在性定理：阿罗不可能性定理。这个定理摧毁了福利经济学的基础，也摧毁了绝对公平信奉者的逻辑基础。

下面我们先简单介绍一下阿罗和德布鲁的成就，然后介绍阿罗德布鲁一般均衡，最后介绍阿罗不可能定理。

简单介绍一下阿罗和德布鲁，这是两个学经济学无法绕开的高峰，张五常就是阿罗的铁粉，因为阿罗可以迅速把一切经济问题变成数学问题。

阿罗（1972年诺贝尔经济学奖）是新古典经济学的开创者之一。除了在一般均衡领域的成就之外，阿罗还在风险决策、组织经济学、信息经济学、福利经济学和政治民主理论方面进行了创造性的工作。

德布鲁（法国数学家和数理经济学家，因为均衡定理证明1983年获诺贝尔经济学奖），他的工作改写了现代数理经济学，他最重要的贡献是与阿罗合作，联名发表了一篇具有划时代意义的文章《竞争性经济中均衡的存在》（1954）。在这篇文章中，运用拓扑学方法，对一般均衡的存在提供了数学证明。他获得诺贝经济学奖的成果一共只有102页：《价值理论：对经济均衡的公理分析》，他开创了一种研究解决问题的先河：德布鲁用集合论和凸性分析来研究均衡问题，彻底摆脱了一般均衡理论主要运用代数和方程的传统，从而彻底解决了亚当·斯密、瓦尔拉斯以来的一般均衡理论只是假设或直觉的逻辑基础（瓦尔拉斯利用代数和方程企图证明一般均衡存在，但是证明被验证是逻辑错误的，因为这种方法本身存在循环假设，无法内在地解决均衡的存在性这一基本问题）。

德布鲁在这102页的证明中，开创了以下概念：资源未被充分利用的度量；帕累托的最优的数学表达（用了数学中的超平面分离定理，在一般意义上建立了价格系统效用最大化配置和帕累托最优配置这两个概念之间的等价性）；相关商品的均衡存在性（一般竞争均衡理论）；用效用函数表示偏好次序关系；总量超额需求函数（效用的需求理论）；经济核算的收敛定理等。

德布鲁的的每一篇文章都给出与经济学核心相关的公理证明，轻松地证明了一个又一个均衡定理。德布鲁114页的《价格理论》奠定了新古典经济学的框架，书中用一般均衡理论讨论了商品、价格、消费等概念的实质意义，还把阿

罗刚拓展的不确定环境中的资源配置问题纳入书中。德布鲁还是纳什均衡（因此获得1994年诺贝经济学奖）的先驱，因为通过讨价还价来决定资源配置，最终也会有一个均衡解，就是德布鲁开始研究的（纳什的成果直观讲就是：在一个复杂经济系统中，每个人根据市场统一的价格进行交易和各人各自讨价还价形成价格的机制是完全不同的。如果价格不同，资源配置的效率自然也就不同。纳什证明：大规模讨价还价最终形成的价格之“核”，是一个不动点，也即存在一般均衡，也即讨价还价是能够实现资源配置的）。

现在德布鲁的成果，已成为微观经济理论的统一构架。他使用的公理化分析方法已成为经济分析的标准形式。

现在介绍阿罗德布鲁模型。

一般均衡理论是经济学理论的核心。一般均衡概念来自于亚当•斯密，也即“看不见的手”：市场会通过价格调整，自动找到供需平衡（直观讲，就是供需曲线一定会相交于一点）。

可是，经济学家们一直没法证明这个均衡点是存在的。

十九世纪末，瓦尔拉斯企图用线性代数来证明这一均衡点存在。但是这个企图失败了。直到1954年，阿罗–德布鲁在《计量经济学》杂志上发表了一篇题为“竞争经济的存在性均衡”论文，提出了阿罗–德布鲁一般均衡模型，用集合论作为基本工具，对经济体制进行了结构抽象，通过一些假设条件，证明了一般均衡的存在，从逻辑上验证了亚当•斯密“看不见的手”的猜想，这是整套新古典经济学的根基。从此，理论经济学被视为科学。阿罗和德布鲁的证明方法简单来说就是：引入一个虚构的市场主体来选择价格体系，从而将给定的经济体系问题转化为一个一般化博弈的均衡存在性问题。

阿罗–德布鲁的证明要用到布劳渥(Brouwer)不动点定理：如果f 是$n+1$维实心球$B^{n+1}=\{x\in R^n+1, |x|\le 1\}$到自身的连续映射（$n=1,2,3,\ldots$），则$f$ 存在一个不动点$x\in B^{n+1}$（即满足$f(x_0)=x_0$ ）。（简单直观说就是任何封闭单位点的连续函数在n维欧几里德空间本身必须有一个不动点）。

这个定理广泛应用于代数方程、微分方程、积分方程等的求解问题，在数学上非常重要，也是微观经济学的基础定理。

这个定理与哥德尔不完全定理，是一切复杂问题解决的大杀器，建议熟悉。例如霍金再《哥德尔与M理论》中就认为，从哥德尔不完全定理出发，建立一个单一的描述宇宙的大统一理论是太可能的。现在人工智能也把哥德尔不完全性定理当成基础定理：根据哥德尔不完全性定理，机器不可能具有人的心智。（哥德尔定理的简单表述是：任意一个包含一阶谓词逻辑与初等数论的形式系统，都存在一个命题，它在这个系统中既不能被证明也不能被否定。或者我们可以这样直观理解：我们永远不能发现一个万能的公理系统能够证明一切数学真理，而不能证明任何谬误。哥德尔不完全性定理的价值是：真与可证是两个概念，可证的一定是真的，但真的不一定可证）。

阿罗-德布鲁构造的模型包括若干假设（完全是描述性的，不具有严格意义）：

假设1：假设存在L种商品，商品的数量为实数，且这里的商品是指最佳划分的商品类，即进一步增加商品类别的划分，由此产生的消费分配并不能增加消费者的满足（保证产出集合凸性假设）。

假设2：假定存在H个消费者，每一个消费者的偏好是一个完备的、连续的、传递的次序，且消费者偏好是非充分满足性和凸性，每一个当事人被赋予拥有每一个厂商的股权（保证消费集合的性假设）。

★集合凸性定义：集合S，对于任意$x, y \in S$，存在$a \in [01]$使得$ax + (1 - a)y \in S$。

假设3：假设存在J个厂商，厂商的生产计划是可行的且能自由决策（这个假设排除了产品的不可分性、规模收益递增和从专业化中获得收益等，且从厂商生产和消费者来说，商品不被加以区分等）。

在上面假设下，阿罗–德布鲁证明：有一组确定的解能够同时满足一般均衡方程组，并且在总量水平上，供给与需求同时均等地决定价格（也即一般均衡状态在完全竞争经济中是可以达到的，并且使之达到均衡状态的价格和产量不是唯一的，只有相对价格的变化才影响消费者、厂商和要素拥有者的决策。如果所有市场在一组价格下处于均衡状态，那么所有这些价格都以同样比例上升或下降后，这些市场仍然处于均衡状态）。

这个定理证明有兴趣的可以去查书，因为涉及太多数学符号，就不介绍了。

阿罗-德布鲁证明一般均衡最主要的假设是：消费与生产集合都是凸集，每个经济主体都拥有一些由其它经济主体计值的资源。

阿罗–德布鲁模型中不需要有固定的生产系数，也不必有一致的利润率，没有股票市场，因为股票不是阿罗-德布鲁商品，模型也不存在企业破产的问题。因为所有的经济主体的生产和消费行为都必须符合预算约束，一旦超过预算，就对其实施无限破产处罚，模型中货币对资源配置没有实质的影响（但是现实中存在货币的理由：交易需求、预防需求、价值贮藏、记价单位等，在阿罗-德布鲁模型中都已经顾及）。

下面介绍阿罗不可能定理。

阿罗不可能性定理的直观表述是：如果众多的社会成员具有不同的偏好，而社会又有多种备选方案，那么在民主的制度下不可能得到令所有的人都满意的结果。

阿罗的课题是：投票选举方式能否保证产生出合乎大多数人意愿的领导者或者说将问题简化为："将每个个体表达的先后次序综合成整个群体的偏好次序"。

阿罗采用数学的公理化方法进行分析，结论是：绝大多数情况下是不可能的。更准确的表达：当至少有三名候选人和两位选民时，不存在满足阿罗公理的选举规则，也即随着候选人和选民的增加，"程序民主"必将越来越远离"实质民主"。

所以阿罗给出了被西方经济学界称为不可思议的定理：如果众多的社会成员具有不同的偏好，而社会又有多种备选方案，那么在民主的制度下不可能得到令所有的人都满意的结果。也即少数服从多数民主原则并不能将个人的偏好汇集成社会的偏好。

简单说明：假如有一个非常民主的群体，或者说是一个希望在民主基础上作出自己的所有决策的社会，对它来说，群体中每一个成员的要求都是同等重要的。一般地，对于最应该做的事情，群体的每一个成员都有自己的偏好。为了决策，就要建立一个公正而一致的程序，能把个体的偏好结合起来，达成某种共识。这就要进一步假设群体中的每一个成员都能够按自己的偏好对所需要的各种选择进行排序，对所有这些排序的汇聚就是群体的排序了。

定理简单说明：假设甲乙丙三人，面对ABC三个备选方案，偏好排序如下：甲$(a > b > c)$；乙$(b > c > a)$；丙$(c > a > b)$（甲$(a > b > c)$代表——甲偏好a胜于b，又偏好b胜于c等等）。

1、若取a, b对决，那么按照偏好次序排列如下：甲$(a > b)$；乙$(b > a)$；丙$(a > b)$；社会次序偏好为$(a > b)$；

2、若取b, c对决，那么按照偏好次序排列如下：甲$(b > c)$；乙$(b > c)$；丙$(c > b)$；社会次序偏好为$(b > c)$；

3、若取ac对决，那么按照偏好次序排列如下：甲(a¿c)；乙$(c > a)$；丙$(c > a)$；社会次序偏好为$(c > a)$

于是得到三个社会偏好次序–$(a > b)$、$(b > c)$、$(c > a)$，其投票结果显示社会偏好有如下事实:社会偏好a胜于b；偏好b胜于cc；偏好c胜于a。显而易见，这种所谓的社会偏好次序包含有内在的矛盾，即社会偏好a胜于c，而又认为a不如c。所以按照投票的大多数规则，不能得出合理的社会偏好次序。

所以依靠简单多数的投票原则，要在各种个人偏好中选择出一个共同一致的顺序，是不可能的。这样，一个合理的公共产品决定只能来自于一个可以胜任的公共权利机关，要想借助于投票过程来达到协调一致的集体选择结果，一般是不可能的。

证明阿罗不可能性定理需要利用May关于完美投票的定理。

May完美投票定义：

★没有弃权票（也即选票定义无限制）；

★两个投票者相互换票, 获胜者不变（选票对称且匿名）；

★投票系统对候选人平等（也即当所有投票者换投另一人, 相应的胜出候选人也会改变。例如简单多数胜出就是一个平等的投票系统）；

★单调性（也即胜出者不会因为得到更多的票而失去胜利，失败者不会因为失去票而得到胜利）。

May定理：如果有奇数个投票人，那么多数胜出投票是唯一符合上述4条规定的投票系统。

May定理的数学形式：设N是一个有限集（全体投票者，假设有2n+1个人），N的全体子集记为$P(N)$，$\{0, 1\}$为两点集, 表示候选人,投票为函数$v : N \to \{0, 1\}$, 记V为v的1原像集, 即投候选人1 的投票人全体, 则V是N的一个子集合，反过来也一样, 即N的任何子集合都可能作为投候选人1的全部投票人。以$|V|$记V中的人数。投票系统为集合函数$f : P(N) \to \{0, 1\}$ (集合函数即集合的子集合类上的函数)

多数胜出投票系统指: 对任意$V \in P(N)$, $|V| > n$ 时, $f(V) = 1$, $|V| \leq n$ 时, $f(V) = 0$

将N排序, 以π 记N的一个位置置换(重新排序), 排好顺序的N的一个A, 以πA记A在π的换位下得到的新的N的子集合，定义域无限制相当于v的定义域是N。假设满足下面三个条件：

★Neutrality：$A \in P(N)$, A^c记A在N中的余集(补集), 则$f(A^c) = 1 - f(A)$。

★Monotone：任意$A, B \in P(N)$, A包含于B中, 则有$f(A) \leq f(B)$。

★anonymous：$f(\pi A) = f(A)$

则有下属结论（May定理）：在有奇数个投票者即$|N| = 2n + 1$时, 多数优胜系统f，是唯一满足上述三条性质的集合函数。

根据上述定义，有两个特殊的投票系统或者集合函数，一个是独裁系统，即除了一个投票人之外，所有的票都被无视，相当于集合函数只在包含这一个人的子集合上非0。另外一个是强加系统，也即存在一个候选者无视所有的投票, 计票输出函数强行赋值给1，也即投票系统函数是平凡的集合函数。

显然根据上述May定理，只需要简单改造，就能证明阿罗不可能定理，例如：

★投票系统仍然为集合函数, 但函数的取值为候选人的一种名次排列(全序)，要求每个投票人给出一种候选人的全序, 投票系统必须给出候选人的全

序，同时候选人平等（也即候选人置换后, 投票系统函数输出的全序不变）；

★传导性(把May假设从Neutrlity换成Consensu) 如果对两个候选人,每个投票人都给出同样的序关系$A > B$, 则投票系统对此二候选人给出同样的序关系$A > B$;

★第三方独立性（即投票系统输出两个候选人的序关系, 独立于其他候选人）；

★非独裁性（即如果有一个投票者(独裁者), 同意$A > B$, 其余投票者都相反, 则投票系统输出中, $A < B$）。

这样假设改进后，极易证明阿罗不可能定理：不存在这样的投票系统, 满足上述四个条件, 还能输出一个全序出来。

阿罗证明的思路：假设群体S上有m个个体成员，群体中出现的各群体种事件构成一个集合X，每个个体对每一事件都有自己的态度，即每个人都对集合X有一个偏好关系>（$i = 1, 2, \ldots, m$）。即可以按自己的偏好为事件排序。

定义群体的偏好为：其中P是一种由每个个体偏好得出群体偏好的规则。按这个规则从个体排序(偏好)得到群体排序(偏好)，而且这个排序符合民主社会的民主决策的各种要求。注意这个排序是自反的，即如果$A > B$，那么，$B < A$；是传递的，即如果$A > B$，$B > C$，则有$A > C$；并且还是完全的，即要么$A > B$，要么$B > A$，二者只有其一而且必有其一。这首先要考察一下民主社会的民主决策的各种要求是什么，阿罗用4个公理表述出这些要求。

公理1：个体可以有任何偏好；而且是民主选择–每个社会成员都可以自由地按自己的偏好进行选择;

公理2：不相干的选择是互相独立的;

公理3：社会价值与个体价值之间有正向关联;

公理4：没有独裁者–不存在能把个体偏好强加给社会的可能。

阿罗证明，满足这4条公理表述的要求的民主决策的规则是不存在的：如果X中的事件个数不小于3，那么就不存在任何遵循原则U，P，I，D的规则(称为社会福利函数)。这表明满足所有一般条件的民主选择要么是强加的，要么就是独裁的结果。

换句话说，阿罗不可能性定理指出，多数规则的一个根本缺陷就是在实际决策中往往导致循环投票。

在得多数票获胜的规则下，每个人均按照他的偏好来投票。不难看出，大多数人是偏好X胜于Y，同样大多数人也是偏好Y胜于Z。按照逻辑上的一致性，这种偏好应当是可以传递的(transitivity)，即大多数人偏好X胜于Z。但实际上，大多数人偏好Z胜于X。因此，以投票的多数规则来确定社会或集体的选择会产生循环的结果。结果，在这些选择方案中，没有一个能够获得多数票而通过，这就是”投票悖论”，它对所有的公共选择问题都是一种固有的难题，所有的公共选择规则都难以避开这两难境地。

那么，能不能设计出一个消除循环投票，做出合理决策的投票方案呢阿罗的结论是:根本不存在一种能保证效率、尊重个人偏好、并且不依赖程序(agenda)的多数规则的投票方案。简单地说，阿罗的不可能定理意味着，在通常情况下，当社会所有成员的偏好为已知时，不可能通过一定的方法从个人偏好次序得出社会偏好次序，不可能通过一定的程序准确地表达社会全体成员的个人偏好或者达到合意的公共决策。

证明过程：

第一步：假设可以找到一个决定群体, 或者独断群体, 这个群体如果一致投$A > B$, 投票系统将输出$A > B$（这个群体类似独裁系统，他把这个群体叫

做quorum）；

第二步：由于候选人的平等性条件，quorum与具体的候选人A, B无关；

第三步：如果投票人子集合S，T都是quorum, 则$S \cap T$也是quorum；

第四步：从全体投票人去掉随便一个人，由于非独裁性条件, 这个集合必然是一个quorum；

结论：所有这些（全体投票者中去掉一个人）的quorum, 相交也应该为quorum, 但是为空集，这与与传导性条件2矛盾。证毕。

这个定理曾经震惊西方，因为证明一个社会不可能有完全的每个个人的自由——否则将导致独裁；一个社会也不可能实现完全的自由经济–否则将导致垄断。也即证明了社会福利曲线不可能存在。

阿罗的不可能定理对福利经济学是毁灭性的。因为福利经济学假设公平是存在的。

所以当年一大批经济学家痛恨阿罗，因为敲掉人家饭碗了，李特尔，萨缪尔森等等领衔的上百篇文章对他的定理进行驳斥。但是阿罗的不可能性定理经受住了所有技术上的批评，其基本理论从来没有受到重大挑战，可以说无懈可击。所以后来这种攻击也就不了了之。毕竟萨缪尔森也是诺贝经济学奖获得者（1970年诺贝经济学奖获得者，第一个获得经济学奖的美国人），而且是数理经济学的先驱人物，是第一个把数学分析应用于经济学研究中的大家，数学也是不错的。不然也证明不了大道定理这种完全超出人直觉的东西。

大道定理简单说就是：在多部门经济体系中存在着多条均衡增长路径,其中增长速度最快的一条称为“大道增长路径”或称为“冯·诺依曼路径”。形象说法就是：当我们要尽快从A点走到B点时，最好的办法就是是先从A点走到高速公路上，然后高速向前，最后才离开高速公路而到达B点。这条高速公路就是最有增长路径。萨缪尔森把制度变量作为约束条件，建立多部门多变量的经济增长模型，利用变分法的的欧拉方程和最优控制的庞特里亚金最大值原理，求出其一阶条件，并解出最优路径，并证明最优路径是收敛的，这就证明了在多条均衡增长路径中，存在一条最优增长路经或最快增长路径。当然假设不同条件下就会有不同的大道定理。

另外萨缪尔森还证明了著名的斯托尔珀—萨缪尔森定理：简单说就是：长期来看，开展国际贸易后，出口产品生产中密集使用的生产要素（也就是本国充裕的生产要素）的报酬会提高，而进口产品生产中密集使用的生产要素（也就是本国稀缺的生产要素）的报酬会下降，而且无论这些生产要素在哪个行业中使用都是如此。

萨缪尔森另外一个定理是：赫克歇尔—俄林—萨缪尔森定理：只要存在产品价格的差异，两国就会继续开展贸易，但最终的结果将是两国两种产品的价格完全相等，而生产要素的价格也完全相等，此时如果其他条件不变，贸易也就停止。也即两国间开展贸易的结果会使两国的生产要素价格最终相等。这个定理也被称为要素价格均等化定理。

当然萨缪尔森还有其他成就，例如福利经济学里面的社会福利函数。

这里顺便瞎扯一下福利经济学。西方经济学大概分为三类，一类是纯粹的理论经济学，大概就是用经济学名词的数学，例如现在基于微分几何，代数拓扑的微观经济学，宏观经济学和计量经济学，不是数学系毕业的，基本看不懂。第二类是应用经济学，主要包括三大领域，例如研究一个国家如何获得经济高速增长的发展经济学，发展的成果如何合理分配以保持持续增长动力的福利经

济学，以及保证经济增长和分配公平的制度经济学，这是西方现在应用经济学三大支柱，不过目前也基本是数学；第三类就是一切沾点经济学边的所谓技术经济学，不过大多数理论经济学家不认为他们是经济学，例如产业经济学，区域经济学，环境经济学，金融经济学，信息经济学，劳动经济学，法律经济学，管理经济学，公共经济学，国际经济学，社会经济学，货币经济学，行为经济学，国防经济学等等（目前在中国挂某某经济学的学科有400多个，不过大多数是名不副实，挂羊头卖狗肉，中国社科院有的自封所谓经济学家有的甚至就是民科，有的所谓经济学大家，其实连弹性（供需曲线的一次导数）和边际（供需曲线的二次导数）这种基础概念都弄不清楚。80年代末期斯坦福，哈佛等等大学的一些经济学教授受286邀请来中国讲课，在中国转了一大圈，最后结论就是：只有中科院系统所里才有真正的符合国际定义的经济学家，其他地方都是自己认为的）。

福利经济学研究的主要内容有：社会经济运行的目标，或称检验社会经济行为好坏的标准；实现社会经济运行目标所需的生产、交换、分配的一般最适度的条件及其政策建议等。

福利经济学从提出至今，大致走过了三个阶段。首先是庇古的福利经济学，有两个基本命题：第一是国民生产纯产值越大，社会经济福利就越大；第二是国民收入分配越是平等，社会经济福利就越大。这个理论就是说：国家经济实力上升必然导致公民福利增加。显然中国现在的事实证明这个理论纯属瞎扯，这个理论早就破产了，甚至被成为脑残理论。

第二个阶段是以帕累托最优状态命名的福利经济学，强调整体经济效率增加对福利的好处：在经济运行时，如对现状进行的改变，使得至少有一个人的福利增进了，这种改变就有利；如果使得至少一个人的福利减少了，这种改变就不利；但是，如果使得一个人的福利增进的同时，而使得另一个人的福利减少，就不能说这种改变一定有利或者不利。然而就是在此时，却达到了帕累托最优状态。可以证明：在完全竞争的条件下，如果消费者追求最大效用满足，生产者追求最多利润获得，生产要素所有者追求最大收入，加上没有经济外部效应，就一定能够达到社会福利最大的帕累托最优状态（按照这个理论，容易证明如果一个人独霸全世界的山河土地，仍会存在着相应的一个帕累托最优状态）。这个理论是为资本家剪羊毛提供心理安慰：我剥削有理，我不剥削你，你连饭都没得吃，我吃肉，你喝汤。这个理论被称为鸡脚杆上刮油有理理论，在明智开发的时代，显然也无法生存了。

第三阶段是以伯格森、萨缪尔森为代表的社会福利函数福利经济学，他们认为：帕累托的最优状态只解决了经济效率问题，没有解决合理分配问题，不同的收入分配会对消费和生产发生不同的影响，而经济效率仅是社会福利最大的必要条件，而合理分配产品收入是社会福利最大的充分条件，只有同时解决效率和公平的问题，才能达到社会福利的唯一最优状态。他们以政治投票与货币投票具有相似特征出发，提出用政治投票的方式构建社会福利函数。

但是阿罗摧毁了这些人的公平梦想：无法以投票的方式产生人人都能接受的唯一社会福利函数。

阿罗不可能定理最本质的理解是：个人私自利益与社会整体利益无论如何必然存在矛盾，不能在满足所有个人私有利益的前提下，逻辑地导出社会整体利益同时也被满足的结论。

阿罗不可能定理其实也摧毁了整个西方经济学关于市场竞争好处的理论基础：西方认为市场经济有好处是假设市场经济必然导致资源配置最优化，因为他们假设个人利益被满足的同时，会自动地导向全社会资源配置的总体优化

（福利经济学第一定理：不管初始资源配置怎样，分散化的竞争市场可以通过个人自利的交易行为达到瓦尔拉斯均衡，而这个均衡一定是帕雷托有效的配置；福利经济学第二定理：每一种具有帕累托效率的资源配置都可以通过市场机制实现。也即可以使效率问题与分配问题相分离。有时收入分配的结果并不尽如人意，但这并不能否认市场的作用）。

所以，下面要进一步介绍的实变函数就是以证明存在性为主，而不是计算解为主的学科。一般说来，现在非数学家是无法获得诺贝尔经济学奖的，因为现在经济学研究的问题已经超出直觉和常识太远了，非数学家，完全无法理解问题本质。现在非数学家研究经济学，基本就是民科路线：望文生义，胡搅蛮缠。目前国内一些老年的所谓经济学大家，基本都是望文生义，胡搅蛮缠的模范人物。

顺便说一句，豆瓣对数学公式不支持，尤其是积分，矩阵，复杂求和等等符号无法表达，所以这篇帖子惨不忍睹，写得实在艰难。错误在所难免，发现请指出来。没有仔细检查，先发出来，大家一起找错误，符号太困难，错误可能有点多。

问答部分（注意：排列不一定是按照时间顺序）

专注一点：忙总，您对德扑AI赌神，怎么看阿尔法狗的围棋是完全信息对称的，而德州扑克是非对称的游戏，AI都能赢，是否证明我们即将进入一个崭新的时代？！

wxmang：不了解。

武当七瞎：忙总雄文，由浅入深，一路娓娓道来，由数学入手落到经济，震聋发聩，开拓思维，提升眼界。能在网上遇见忙总这样学通数理又能经世致用的老师，虽然相逢恨晚，仍是三生有幸。结尾的阿罗不可能定理看了好幻灭，既证明了程序民主不可能导致实质民主，又证明了共产主义福利经济学的不可能，那我们终究要追求什么社会目标了？迷茫了……

wxmang：一般说来，人类社会追求的方向，由统治阶层决定，统治阶层选择对他们最有力的方向。例如中国古代，任何统治阶层选择的都是：国泰民安，民富国强，因为国泰民安就稳定，便于延长统治期限；民富就便于剪羊毛，国强，就便于抵抗外来势力颠覆，维持自己千秋万代的人上人地位。所以不管真相或真理是什么，我们都只能接受统治阶层的选择，普天之下莫非王土，率土之滨莫非王臣。

捕风胖子：忙总功力好深厚。。。。说起线性无关，忽然想起李炯生老师讲的：“所以内，线性相关就是注水猪肉，线性无关就是没有注水”，令人印象深刻。

wxmang：你是科大数学系的？听李炯生老师的线性代数就像飞一样，这么好的老师现在找不到了。科大学生基础好，微积分和线性代数基础扎实占至少一半功劳。

猬心：能否请忙总从数学的角度点评一下产业经济学SCP模型的有效性，东北财大和人大有一堆家伙就靠这个吃饭。

wxmang：可以当成游戏，但是没有用。因为产业政策制定，某种程度是各方利益博弈的结果，而不是模型算出来的最优路径或最优方案。举例来讲，以前搞的高端制造业（现在叫智能制造或制造4.0）产业政策，就是央企相关的几大家族自己掏钱找一批专家制定对自己企业最有利的产业政策草案，然后拿到国家发改委会议上吵架，争预算，争政策扶持，然后发改委根据几家方案，汇总整一个分果果的方案，报国务院，国务院再平衡各省的利益，就下达执行了。我搞过几个产业优化模型，得到的结果，不过是为企业争取一点预算和政策扶持而已。产业经济学模型这玩意其实没什么用，不用它，用投入产出模型，我也能吵赢。产业资源配置中，投入产出模型是非常深刻的，一般无人可以用任何其他模型驳倒。如果你熟悉投入产出模型，在发改委系统就无敌了。在全国交通规划中，我们也用投入产出模型通杀。

猬心：谢谢忙总的真相，有人整理过八十年代的会议纪要，那时候大多是协商（吵架）制定当年的发展目标和目标。

雪泥飞鸿：投入产出模型有没有什么书可以推荐？

wxmang：以前推荐过，系统所陈锡康（他是投入产出发明者列昂节夫的学生）的《投入产出技术》，科学出版社

塔里木河：感谢忙总从更高层次对数学的介绍。很想帮忙总安装一个latex editor。可视化界面很方便。这样忙总就可以比较轻松地写公式了。

wxmang：我能写公式，主要是豆瓣不显示数学公式。

塔里木河2017-02-06 11:42:35 我想的是把忙总发出来的文本形式的latex公式拷贝到editor中，就可以了。不过忙总写的过程似乎要添不少麻烦。

laokay：请问忙总，均衡、优化和突变这三类问题，用数学语言如何描述？

wxmang：这本身就是数学语言了。想进一步了解背景，建议看维纳的《控制论》，这是讨论系统均衡的基础模型的；庞特里亚金的《最优过程的数学理论》，这是讨论优化的基本问题的；托姆的《突变论》这是讨论突变模型的。

far8008：感谢忙总大作，大学没有学数学真遗憾，看来数学学好了，是可以看清很多问题的。

wxmang：不一定，主要还是看思维方式。毛主席的数学就很差，照样洞察力惊人。

wxmang：当年我刚到系统所时，喜欢跟老师们聊天（也都是科大数学系校友），讲自己的看到的和感受，讲一大堆后，几乎所有老师的必然问题是：你想说的本质问题是什么（或者你的问题本质特征或本征是什么）。对数学家来讲，所有问题都应该是抽象成本质的问题。他们认为，人类只能解决三类问题：均衡、优化和突变，离开这三类问题的问题，都无讨论价值。实际上系统所的研究室设置也是按照解决这三类问题设置的，例如控制研究室主要解决均衡问题（稳定性，可控性等等）（这个研究室最著名的成就就是红旗系列地空导弹控制系统的理论模型和算法，反导系统控制系统的理论模型和算法，核武器反应堆控制系统的理论模型和算法等等）；运筹学研究室主要解决优化问题（资源最有配置，网络分工优化，能力调度等等）（这个研究室最著名的成果有全国铁路调度系统，国家资源配置系统，指挥自动化和电子对抗系统等

等）；基础研究室主要解决突变问题（例如耗散结构，自组织等等）（这个研究室的成果除了一些复杂系统基础成果外，就是核潜艇通信系统，卫星通信系统，卫星侦察系统的算法和模型等等）。

slyypp：(°-°)…这些研究成果应该公开渠道找不到吧，失落，要有发表的论文找来看看就好了

wxmang：连获奖名单上都不能出现课题的名字，只有编号，有的课题，至今名称都是保密的。

青紫色的脸颊：流口水啊，想去系统所念研究生。

wxmang：现在的系统所已经没有军工项目了，1988年以后，大裁军后，大量项目下马，搞军工项目的，一部分去了总参2部，3部，海军装备论证部，空军装备部等等，一部分改行搞民用项目，例如项目管理，企业管理，一部分出国混饭吃。现在这批人大多数人不是退休就是去世了。

fightp：“8、级数”下，“（2）函数级数”和“（4）泰勒级数”中间少了（3）。不知道内容有没有漏掉

wxmang：内容没少，分段少了一个小标题（3）、函数级数的收敛判别，已经补上了。

司马错：忙总，(1)如果动点P是在一个三维坐标空间(r, s, t)中，则函数应是三元的：$x = \phi(r, s, t), y = \psi(r, s, t)$,雅可比矩阵则是：您此处的矩阵就直接略过不写，是因为豆瓣网站的缘故吗。

wxmang：豆瓣不显示矩阵，我写了，被豆瓣吃掉了。

欧拉陀螺：读完忙总的几篇数学大作，对数学领域几个部分有了粗略理解。有两个问题想请教忙总：1范畴论在抽象数学领域是什么地位，是不是代数结构更深一级的抽象？2我看过一种论调抽象数学和布尔巴基派在法语教材中更突出，据说是为了增强法兰西民族自信心，后来布尔巴基派将这一理念扩展到中小学结果铩羽。抽象数学是不是只是少数人才能驾驭，外行人学这个，最终会不会只是一种事倍功半的高级玩具？

wxmang：我没学过范畴论，不过听搞代数拓扑的人讲（范畴论是代数拓扑搞出来的），就是升级版集合论，是一种语言系统，本身并无什么拿得出手的成就，没有发现或证明任何一个超出直觉和常识的定理。学数学，如果不是想成为大数学家，如果只是靠数学混饭吃不需要什么天赋。以我的经历来看，我没什么数学天赋，但是在学习过程中，并未碰到什么太大困难，虽然水平无法与天才相提并论，但是理解大多数知识毫无问题。而且我的基础很差，上大学学习数学前，一共只上过5年学，上初中时甚至不会除法。所以学习数学无难度。法国数学整体水平大大高于我国，目前仍然名列世界前5名水平（美国，德国，俄罗斯，法国，英国），据在法国知名大学当数学教授的同学讲，法国几个好的理工科大学的数学水平高于中国科大整体水平。我国数学水平不如日本以色列波兰匈牙利等等。2010年38岁的越南数学家吴宝珠因证明了朗兰兹纲领中自守形式中的基本引理而获得国际数学界的最高奖菲尔茨奖震动中国，因为这是土生土长的越南人获奖。土生土长的中国人目前连提名都做不到。目前获得提名或获奖的中国人，其实都是海外华人，与中国数学界没任何关系。我们杀鸡取卵，急功近利的科研体制严重掩杀人才成长。

winternight39：十几亿人的国家连以色列波兰匈牙利这种地级市国家都不如？

wxmang：波兰数学曾经很伟大，华沙与哥廷根曾经是世界两大数学中心之一，华沙学派曾经的地位不次于布尔巴基学派，随便瓣手指，都能数出十多个一流的波兰数学家，例如泛函分析奠基人巴纳赫就是其中代表（泛函分析另外以为创始人施坦因豪斯也是波兰人），波兰数学家在拓扑，集合论，数理逻辑，抽象代数，微分几何等等都有很高成就。匈牙利也类似。现在稍微差点，是因为美国人把一流数学家挖走了，例如国际数学家大会名誉主席奥尔里奇就是美国人挖走的，这是泛函分析和拓扑学的大腕，Orlicz空间、Orlicz－Pettis定理、Mazur-Orlicz定理等等学过都知道。

纸飞机：法国是第一梯队的，略强于德国，高铁非常方便法国和德国数学家进行交流。很多俄罗斯数学家在法国有合作项目，甚至一年中有多半年在法国教课的。天朝现在每年能去巴黎高师读数学也就是可以数出来最好的那几个人，法俄数学培养上基础扎实，比美国好得多。日本很强，第二梯队前列，至于天朝不算海外人才的话，能混到第三梯队就不错了。

wxmang：中科院数学所自己认为中国数学在世界排名在20名开外。美国数学强主要靠挖人。

土豆烧牛肉：关于阿罗不可能定理，其实宗教、文化、法律这些东西都是用来塑造一个群体的偏好，以使得做出关系群体长期利益的决策成为可能。中国几次大乱世后的稳定期可作为例子。

wxmang：中国历史上老百姓就没有过投票权，李自成张献忠那种投票权不能算，那是玩命权。我们传统上，一直是皇帝+精英阶层独裁。

我是你大爷：那以后也会是精英独裁吗？

wxmang：现在难道不是吗？早就是了，1978年以后就是了。

宁静以致远：德国的恩格尼玛密码机据说最早就是被波兰人破译出来的，我觉得这应该和波兰的数学家群英荟萃不无关系吧

wxmang：恩格尼玛密码破译是被波兰的三剑客：杰尔兹·罗佐基，马里安·雷耶夫斯基，亨里克·佐加尔斯基联合破译的，这三个都是密码界的天才，后两位是后来的抽象代数大家，前一位在1942年就阵亡了。

折花落燕：佩服忙总，虽然没有做数学研究，这么多年过去还能记得这么清楚，并且把各种理论的出发点和本质讲的这么清楚。我大学毕业不到4年，已经忘光了微积分的基本知识了，实在惭愧。回想起来，我本科也是国内不错的学校了，但在高等数学的教育还是有很多不足，太偏重于计算技巧而忽略数学思想的培养，太偏重于考试而忽视对问题价值的认识。当时太年轻，也没有其他人能指导自己，导致自己对数学的价值没有清楚的认识，数学是考过就忘，后来才发现数学的巨大魅力和价值，真是悔之晚矣。

wxmang：我的观点是，如果学过就忘掉，何必浪费时间去学习？这不是浪费生命吗？人生不需要学习太多东西，只需要把学过的学精，会用，终身不忘，即可成功。

回车1234：芒总新春愉快！请教芒总一个小问题："阿罗不可能定理最本质的理解是：个人私自利益与社会整体利益无论如何必然存在矛盾，不能在满足所有个人私有利益的前提下，逻辑地导出社会整体利益同时也被满足的结论。"这句话可否理解成"小河有水大河满"是不存在的？反过来"大河有水

小河满”是否也不存在？

wxmang：你说的问题不存在，因为你说的是分配问题，分配问题是政策问题，由统治者决定，想分你多少就是多少，税收和基准工资就是分配工具。实在不行还有通货膨胀率可以调节，通过高通胀来杀富济贫。阿罗定理只能得到这样的推论：个体理性不能得到整体理性；整体理性代表不了个体理性。换就话说就是：整体效率最优，一定是牺牲部分个体利益实现的；部分个体利益最优，一定是占大家的便宜实现。

泰坦之刃：那么既然福利经济学仆街了，取代它的新理论是什么？

wxmang：阿罗不可能定理后，福利经济学做了两次大的根本修补，一次是阿马弟亚·森(AmartyaKumarSen）用最小选择函数代替萨缪尔森的社会福利函数后，证明的最小自由定理：也即人类可以牺牲个人的自由换取社会福利公平（不过这个定理在西方强调民主自由意识形态内引起混乱，也即赞美了独裁的效率和公平是可以均衡的）；第二次是2004年艾利亚斯证明的定理：如果有多于两个可供选择的社会状态，那么，任何社会集结算子，只要满足”偏好逆转”假设和”弱帕累托”假设，就必定是独裁的。这个定理本质就是任何公平和效率的均衡，代价都是独裁。无法避免。民主只能带来低效率和不公平的失衡。可以从数学上证明，阿罗的不可能性定理、森的最小自由定理，缪勒和塞特斯维特的帕累托效率兼容的不可能性定理等等，均可视为艾利亚斯一般不可能性定理的特例。说亿i，现在西方福利经济学不再讨论公平，民主之类问题，而是讨论外部资源对福利体系的支撑，次优理论，相对福利学说（有限福利或不养懒汉福利或部分人不能享受福利）、公平和效率交替学说、宏观福利理论等等等，这些“新”理论一方面企图说明，现代西方国家可以通过政府干预调节价格和产量，实现资源的合理配置；另一方面企图说明，现代西方国家的分配制度虽不合理，但是如果加以改变，则可能更不合理，一切人为的改善分配状况和增进福利的措施都是无效的。无耻吧？无耻至极吧？

Cousteau：“数学中有哪些明明是暴力破解还给人美感的证明？”，这个回答刚好是arrow定理的简化证明。

https://www.zhihu.com/question/24239308/answer/112972302

土豆烧牛肉：只能说现在的经济学都是为体制服务的。还是忍不住问一下，数学能证明共产主义社会实现的可能性吗？

wxmang：不能，因为经济学基础假设（或公理）：资源有限。共产主义公理假设是：要求物质极大丰富，可以按需分配。本身就矛盾。不过从现实世界来看，显然经济学公理是正确的。

大脚丫：高通胀应该是劫贫济富吧？一般来说通胀直接增加生活成本，但是对产业资本来说销售和利润增加是好事啊，富人才不怕涨价呢。忙总是不是笔误了？

wxmang：没有笔误，我在以前介绍通胀的帖子里反复说过，你没仔细看。通胀主要是让存款贬值和固定资产贬值（尤其是房地产），穷人是没什么存款和固定资产的。穷人靠工资生存，工资会随通胀率水涨船高，所以对穷人影响不大。蒋介石在48年以后搞恶性通胀，杀富济贫，结果导致丢掉政权，反对他的不是穷人，而是被他搞穷的富人。建议看看我写的金圆券的帖子。

花在田：在美国穷人负债高，通涨是抢富人的钱，在中国富人负债高，穷人才储畜，通涨是抢穷人的钱

wxmang：中国个人存款56万亿，法人存款126万亿，你觉得谁多？如果通胀能够剪穷人羊毛，你觉得周小川顶得住，央行会手软？

大脚丫：我还是受自己头脑中原有的框框限定了，你的那些帖子其实我都看过，但是限于知识储备，可能都没有看进去。通胀劫富济贫和我的直觉相反，金圆券也是说民不聊生啊，没注意到是富人吃亏更多。股市上也是这样说，通胀时资产价格要增加。美国人也是号称二十年没涨岗位工资了。这些社会现象在我看来都是穷人受苦，现在哪里有年年涨工资的穷人呢？不过我这些都是纯粹个人感受，谈不上什么理性。

自然的风吹来：其实大家感受都一样，中国通胀是反过来的。中国底层的工资真不一定比物价涨的快，就像有网友说了，中国富人负债多，穷人反而把打工的血汗钱拿去存款，通胀成了劫贫济富，通胀造成资产价格上涨，拉大了贫富差距，我想这些大家感受都一样的。这点忙总肯定错了

wxmang：这是经济学常识问题，你觉得我连常识都会错？好歹我也是第一届国家反通货膨胀课题组的成员之一。如果通胀真的主要是剪穷人羊毛，周小川顶得住？央行会手软？面对复杂问题，直觉往往是错误的。

[已注销]：你的观点：通胀主要是让存款贬值和固定资产贬值。他的观点：中国富人负债多，穷人反而把打工的血汗钱拿去存款。我觉得他不是不认可经济学常识，而是不同意你对现实的认识。

云风：对什么是富人，什么是穷人认识有偏差。56万亿和126万亿比，而不是自己和身边的富人比。身边负债的富人未必是真的富人。

[已注销]：对，就是这个意思，普通人对富人没概念。不过对于这56和126万亿，倒是不知道其中分别有多少是贷款。

西穆坞:房产应该是抗通胀比较好的资产吧？

wxmang：这个不能一概而论，当然只能考虑一线城市房产了，三四线就不要指望保值增值了。就算是北上广深，也还要考虑地段（非成熟地段或配套不齐全的地段，即使是北京也无保值增值价值），然后是类型（商业地产还是驻宅，住宅类型，面积，小区档次，物业，质量等等），三是购买价格（高价接盘就不存在保值增值了），四是贷款比例（全款无贷有最强的保值增值能力，杠杆太高，就没有保值增值能力了，因为贷款利率也会随通胀水涨船高，吃掉利润）。如果这些条件都好，有保值增值空间，

Cousteau：私以为波兰匈牙利等欧洲国家的数学优势和修道院经院哲学的教育传统有很大关系，必修课就是不同的神学抽象元素间颠来倒去。

wxmang：不是，巴纳赫说过，波兰数学繁荣，主要是他们国家太穷，没钱给科学家搞物理和化学，只能搞不要什么钱的数学。

Cousteau：那么单纯说经院环境对抽象思维的影响呢？比如荷兰小国的布劳威尔的数学直觉主义？

wxmang：荷兰数学本身也是不错的，基础很好，惠更斯就是荷兰人，不过他的大数学家不多，近代的数学家里面，范德瓦尔登，布劳维尔大概算最好的了。

Cousteau：忙总宕开问题了... 您提到波兰数学，波兰出身的Mandelbrot做分形是有构造函数的。正文中说“在数学中，证明解的存在性比如何计算解本身要重要得多”单纯就数学问题或数学证明来说，就您个人口味而言，请教下是比较欣赏给出具体解的构造或算法的证明，还是倾向于更加单纯（极可能富有“哲理”或“开拓性”）的存在性证明呢？

wxmang：我的数学水平尚不能判断这种问题。

香辣蟹：忙叔觉不觉得，解释通胀的时候，就像你曾经说的：你叫不醒装睡

的人？以前我从来没想到房子及其衍生的财富，以及由此而来的社会失衡能如此深刻的改变社会氛围和文化。从08年开始这几波羊毛剪的也许都快媲美刚解放的土地革命对社会的影响了。

奔理指归：金融能够割韭菜，根本原因是人类大脑中特定部位具有的生理特性，能够对外界刺激产生恐惧和贪婪，形成正反馈，在这个过程，绝大多数人根本无法控制，赌博上瘾原因也是如此，这样的生理性BUG，让做局的人屡试不爽，沉醉其中的人，会按照他们期待的结果去寻找一切具有说服力的理由，去说服自己，也去说服别人（索罗斯金融炼金术的反身性就是这个意思），但是其实他们内心是极其恐惧的，这就是所谓的，你叫不醒装睡的人，其中经典的一个理由就是通胀会助推房价，债务会因为通胀稀释，他们其实不知道他们这几年领的高薪其实也是资产泡沫支撑的，泡沫一破，直接失业了，还有毛线的工资领。浑然不知也是种幸福，其实装睡也挺好的

wxmang：其实这个在西方教科书中早就说过：革命是直接剥夺财富保证社会公平，政策是间接剥夺财富保证社会效率和公平的均衡（财政政策，尤其是税收政策是二次分配工具，保证大家有饭吃；货币政策，尤其是利率政策，是调节贫富悬殊工具，保证大家不造反）。所以政策比革命效果好，政策强调了发展，而单纯革命而不强调效率，会导致发展萎缩，导致社会进步衰退。