



普通高等教育“九五”国家级重点教材

大气科学中的数学物理问题

黄思训 伍荣生 编著

气象出版社

普通高等教育“九五”国家级重点教材

大气科学中的数学物理问题

黄思训 伍荣生 编著

气象出版社

内 容 简 介

本书根据十多年来教学实践和科学研究的总结,以国内外大气科学发展趋势为思路,吸取近年来最新研究成果,并考虑到气象工作者及科研人员实际业务发展与提高的需要。本书以变分为出发点,介绍变分原理及其算法的数学知识,介绍 Hamilton 力学及最优控制,并给出这些理论在大气科学和流体力学中的某些应用,特别是在稳定性方面的应用。为了读者阅读方便,我们增加了必要的数学基础知识,另外根据实际需要,简明扼要地介绍了近代数学的一些基本概念(例如流形、广义函数、反问题的正则化思想、变分同化技术等)。全书深入浅出,物理概念清晰,并注重理论与实际相结合。

本书经教育部审定为普通高等教育“九五”国家级重点教材,可作为高等院校大气科学专业及相关专业的硕士生和博士生教材,也可作为气象、海洋、航空、农林、水利、环境等部门的科研人员和业务人员的参考用书。

图书在版编目(CIP)数据

大气科学中的数学物理问题/黄思训,伍荣生编著. —北京:气象出版社,2001.3
ISBN 7-5029-3103-1

I. 大… II. ①黄…②伍… III. 大气科学-数学物理方法 IV. P40

中国版本图书馆 CIP 数据核字(2001)第 10777 号

大气科学中的数学物理问题

黄思训 伍荣生 编著

责任编辑:陶国庆 终 审:周诗健

封面设计:曾金星 责任技编:谷 青 责任校对:谷 青

出版发行:气象出版社

出版社地址:北京海淀中关村南大街 46 号

出版社电话:68407112

电子邮箱:cmp@rays.cma.gov.cn

邮政编码:100081

传真号码:62176428

经 销:新华书店总店北京发行所
印 刷:北京昌平环球印刷厂

开 本:787mm×960mm 1/16
印 张:34.25
字 数:686 千字
定 价:50.00 元

版 次:2001 年 3 月第 1 版
印 次:2001 年 3 月第 1 次印刷
印 数:0001~2000

版权所有 侵权必究

前 言

大气科学中数学物理问题,主要指变分问题及其应用。本书是作者根据上述问题,通过对南京大学大气科学系四届博士生的讲授,其间又不断征求同行及有关专家的建议,并听取学生们的意见,在此基础上作了重大修订改写而成的。我们的目的是让博士生和硕士研究生对这一领域内容与发展有一个较全面和深刻的理解,使他们在学習过程中能打下较深厚的数学物理基础。实践证明,这门课程的开设,对于他们掌握近代数学知识,并利用这些知识开展大气科学的理论研究,具有重大的影响。

大气科学已越来越向精确化和量化的方向发展,数学物理在大气科学中的应用也越来越深入。当前极需要一本介绍较深层次的数学物理知识及其在大气科学中应用的书籍,为了满足这方面的需要,我们编写了本书。本书以变分问题及大气运动稳定性为主线,全面讲述变分原理、计算方法和稳定性理论。全书自成系统,章与章之间既相对独立,又互相联系,同时还兼顾了在大气科学中的应用,如四维同化技术、反问题的正则化方法及在卫星遥感反演中的应用、变分与稳定性等。作者考虑到一般气象工作者的数学基础,还特别增加了第一章的数学预备知识。在编写过程中,作者既注意了全书的可读性,力图以普及的形式,深入浅出地介绍近代数学方面的有关知识,又保证了科学体系的完整性。为了阅读方便,作者在目录和正文的某些章节上标注了“*”号,表

示初学者可以不必阅读这些章节的内容。

在完成本书的过程中,受到解放军理工大学气象学院、南京大学大气科学系的领导的热情支持和协助,南京大学谈哲敏教授、潘益农副教授对本书提出了不少建设性的建议,解放军理工大学气象学院项杰博士仔细审阅和修改了第七章的内容,王兴宝博士为本书提供了不少有价值的材料,在此表示深深的谢意。

本书的审稿会于2000年4月在南京大学召开,南京大学赵鸣教授、谈哲敏教授、潘益农副教授、解放军理工大学吕美仲教授、张铭教授、气象出版社陶国庆副编审等参加了审稿会,他们认真审阅了本书初稿,并提出了许多建设性的意见,在此向他们表示由衷的感谢。本书得以很快出版,得到气象出版社的大力支持,在此表示感谢。同时还要感谢倪立琴、韩振华两位女士,为本书的出版录入文稿,付出了辛勤的劳动。

由于作者水平有限,加之时间仓促,书中不妥甚至错误之处一定不少,恳请读者批评指正,并期待读者建议和意见。

作 者

2000年5月于南京

目 录

前言

引言 (1)

第一章 数学预备知识介绍 (13)

§ 1.1 常微分方程(组)的解法 (13)

§ 1.2 常微分方程的稳定性及定性理论 (26)

§ 1.3 偏微分方程的特征理论 (46)

§ 1.4 摄动的基本概念和问题 (57)

§ 1.5 奇异摄动方法介绍 (67)

§ 1.6* 积分方程的理论与解法 (81)

参考文献..... (100)

第二章 数学物理中的变分问题 (102)

§ 2.1 函数的极值问题 (102)

§ 2.2 泛函极值与变分问题 (110)

§ 2.3 力学中的变分原理 (126)

§ 2.4* 可变区域上泛函的变分与 Noether 定理 (134)

§ 2.5* 流体力学中的变分原理 (142)

§ 2.6 平均变分原理和波动传播及发展 (148)

参考文献..... (155)

第三章 一阶非线性方程与 Hamilton 力学 (157)

§ 3.1 一阶拟线性方程的理论 (158)

§ 3.2 一阶非线性方程及其解法 (165)

§ 3.3 全积分与求全积分的方法 (174)

§ 3.4 Legendre 变换及其在力学上的应用 (179)

§ 3.5 Hamilton 正则方程 (185)

§ 3.6* H-J 方程及 H-J 理论 (192)

§ 3.7 变分与最优控制 (204)

§ 3.8* Hamilton 力学的几何原理 (220)

参考文献..... (235)

第四章 稳定性与变分	(236)
§ 4.1 流体稳定性的一些基本概念	(236)
§ 4.2 平面平行流的边值问题特征值估计与变化	(245)
§ 4.3* Couette 流及其稳定性	(253)
§ 4.4 无粘性流体中不稳定理论	(263)
§ 4.5 流体稳定性的能量方法	(274)
§ 4.6 平均变分法与波的稳定性	(283)
参考文献	(294)
第五章 变分原理的数学基础	(296)
§ 5.1 线性空间与对偶空间	(296)
§ 5.2 欧氏空间	(305)
§ 5.3 欧氏空间中微分流形与 Stokes 定理	(316)
§ 5.4 泛函分析中一些基本知识	(326)
§ 5.5* 广义函数初步	(340)
§ 5.6* Sobolev 空间大意	(351)
§ 5.7* 非线性泛函分析介绍	(360)
参考文献	(369)
第六章 变分问题及其算法	(371)
§ 6.1 正算子与二次泛函的极小问题	(371)
§ 6.2* 算子形式的 Euler 定理、互补变分原理及变分的逆问题	(380)
§ 6.3 Ritz 与 Galerkin 方法	(391)
§ 6.4 有限元方法	(401)
§ 6.5 数学物理中反问题正则化方法思想	(411)
§ 6.6* Tikhonov 正则化与变分问题	(422)
§ 6.7* 最大熵方法及其应用	(439)
§ 6.8 R^n 中变分原理及算法	(451)
§ 6.9 四维同化、变分同化及伴随方法	(460)
参考文献	(473)
第七章 流体及大气运动的非线性稳定性	(476)
§ 7.1 流体运动稳定性的弱非线性理论	(476)
§ 7.2 流体运动非线性稳定性的 Arnol'd 方法	(487)
§ 7.3* 多层准地转流的非线性稳定性	(501)
§ 7.4 两层准地转流的非线性稳定性	(511)

§ 7.5* Phillips 模式的非线性斜压不稳定的饱和问题——上界估计	(520)
§ 7.6* Phillips 模式的非线性不稳定的饱和问题——下界估计	(529)
参考文献.....	(535)
附录：常用数学符号及代表意义	(537)

注：目录中注 * 的内容初学者可以不必阅读

引 言

一、变分与各学科之间的关系

Euler 指出:“因为宇宙的结构是最完善的而且是明智的上帝的创造,因此,如果宇宙里没有某种极大或极小的法则,那就根本不会发生任何事情”。Euler 以特有的数学家的透彻、犀利的目光,指出了极大或极小的法则在宇宙中的特殊地位。

极值问题是变分的基本问题。变分理论的发展与力学等其他自然科学的发展密切相关,相互促进。与变分有关的问题很多,例如:数学物理问题(力学,天体力学,几何光学,弹性理论,流变学,相对论等)、几何问题(短程线,极小曲面等)、控制技术问题(一般的最优控制,工业生产自动化等)以及经济问题(迁移问题,最优储存问题等)。另外,在化学、地球物理、CT 技术中,最优决定未知参数的问题;在概率论中,随机过程的最优控制问题等也与变分问题有关。

下面从几个角度来阐明变分与各学科之间的相互关系。

1. 从力学的发展来看变分问题

自 1696 年 J. Bernoulli 提出变分问题以来,变分问题与力学问题相互渗透、相互发展、相互促进。1788 年 J. L. Lagrange 出版《分析力学》一书,书中充分体现了经典变分与力学相结合的特点,被 C. G. Jacobi 称为“分析力学之母”。Lagrange 提出的积分的最大最小方法及推广的最小作用量原理,十分有效地解决各种动力学问题。1834 年, W. R. Hamilton 把最小作用量原理推进到一个崭新的阶段,提出了稳定作用量原理;1829 年, C. F. Gauss 提出力学中最小约束原理,这是最早微分变分原理,这些原理无论在理论上或在应用上都极有价值;1941 年, N. G. Chetaev 把上述各种原理推广到连续介质中去。20 世纪以来,变分原理在力学各分支中找到更为广泛且有成效的应用,尤其是固体力学中变分原理的理论研究与实际应用始终处于前列。例如钱令希(1950)的余能原理, Reissner (1950)的广义变分原理、胡海昌(1954)、龔津(1955)的广义变分原理、钱伟长(1964 年)的 Lagrange 乘子法等对力学中的变分原理的发展起到推动作用,意义深远^[1~5]。

随着变分问题的发展,不变性变分问题也相继产生。1918 年, E. Noether 概括了力学研究中潜在的不变性,对称性,得到了两个不变性原理,称为 Noether 定理。由于 Noether 定理与守恒律相关,所以 Noether 定理引起数学、物理、力学各界对于不变量理论的研究的关注。1962 年 W. Guenther 应用 Noether 定理研究线弹性静力学的守恒律问题。1968 年, J. R. Rice 提出 J 积分的概念,促进了断裂力学的深入研究。实际上, J 积分是一个守恒定律。此后,不少学者在这方面做了大量的工作,找到一些过去不曾

发现的新的守恒律,对探讨力学问题无疑是十分重要的。

2. 流体与大气中的稳定性与变分的关系

运动稳定性问题是流体及大气动力学的经典问题之一。在运动稳定性中,线性稳定性理论发展最为成熟。19世纪末期,线性稳定性的标准模方法开始广泛应用,最著名的是不可压缩流体的线性稳定性判据,即 Rayleigh 定理;1950年, Fjørtoft^[6]改进了 Rayleigh 定理。1949年,郭晓岚考虑了旋转等因素,得到了大气运动的正压不稳定判据^[7]; Charney^[8]和 Eady^[9]研究了大气运动的斜压不稳定性,得到了斜压不稳定性判据。但线性稳定性理论只适用于小扰动情形(因而忽略了扰动与基流的相互作用),于是摄动方法被应用于稳定性问题的研究(Stuart^[10]),弱非线性理论应运而生;1970年, Pedlosky^[11]把弱非线性理论应用于大气运动的稳定性的研究。但弱非线性稳定性理论仍有自身的不足,国内学者周恒等^[12~14]在这方面作了更深入的研究。1965年, Arnol'd^[15,16]首次利用变分原理与先验估计相结合的方法,证明了平面无粘不可压定常流的非线性稳定性。因此,关于稳定性问题的研究进入了一个新的阶段。之后, Arnol'd 的方法(又称为能量-Casimir 方法)得到进一步的推广和发展。如, Blumen^[17], Holm et al^[18], Abarbanel^[19], McIntyre et al^[20]等等。Salmon^[21]对 Hamilton 力学作了系统的介绍,并着重指出了在运动稳定性问题研究中的应用。在国内,关于流体与大气中的非线性稳定性问题的研究也取得了很大的进展。如,曾庆存^[22]、穆穆等^[23,24]结合大气和海洋中的一些重要模式,利用变分原理研究了非线性稳定性问题,得到了很有意义的理论结果。

3. 变分与最优控制、四维同化技术

自 Wiener^[25]1947年创立控制论以来,控制论已渗透到工程、生物、经济、社会等各个领域,逐步形成了一套完整的控制理论体系。所谓最优控制是指:根据已建立的控制对象的时域或频域的数学模型,选择一个容许的控制律,使得被控对象按预定要求运行,并使给定的某一性能指标达到最优值。从数学观点来看,最优控制理论是求解一类带有约束条件(常微分方程,偏微分方程,随机微分方程等)的泛函极值问题。这类问题与变分有着密切的关系。利用经典变分思想可以解决容许控制属于某个开集的一类最优控制问题,但在实际问题中,容许控制往往属于闭集,对于这类控制,经典变分变得无能为力,而取而代之的是近代变分理论,近代变分理论最常用的两种方法是 Bellman 的动态规划与 Pontrjagin 的最大原则。

动态规划是由美国数学家 Bellman^[26]在 50 年代为了优化多级决策问题的算法而提出的,这种方法与原苏联数学家 Pontrjagin^[27]最大原则和 Kalman 滤波理论被称为现代最优控制中的三个里程碑。Bellman 依据最优性原理,发展了变分学中的 Hamilton Jacobi 理论,解决了控制有闭集约束的变分问题,有关方面内容可以参考专著^[28,29]。

在以往的最优控制问题中,认为控制系统不受随机干扰的影响,故一般称为确定性系统的最优控制,但在一个系统中,或多或少受到随机干扰的作用,研究受到随机干扰

作用的系统称为随机控制问题, Wiener 和 Kolomogorov 开创的滤波和预报理论是随机控制理论的里程碑, 有关方面内容可以参考专著^[30]。

70 年代以后, 偏微分方程的最优控制有了很大发展^[31], 并把偏微分方程的最优控制应用到大气数值预报中去, 形成目前十分热门的四维同化技术^[32~34]。众所周知, 大气状态在四维空间中变化着, 为了精确地描述大气的运动状态, 人们对大气的动力及物理过程进行不断地研究与改进, 利用各种方法及方案, 提高业务模式的精度和分辨率。所谓四维同化问题, 就是利用一切有效的信息, 去估计出数值预报所需要某时刻的大气状态值。大气数值模式可表示成如下的非线性发展方程的初值问题:

$$\begin{cases} \frac{dx(t)}{dt} = F(x, t) & t \in (0, T) \\ x(t)|_{t=0} = u \end{cases} \quad (1)$$

其中 u 表示大气数值模式积分的初值, 模式(1)可以由初值 u 唯一地确定解 $x = x(t, u) \equiv x(t)$, 记 $H(x(t), t)$ 为时刻 t , $x(t)$ 与观测数据 x_{obs} 之间的距离, 由此在 0 到 T 时间内 $x(t)$ 与 x_{obs} 所产生的总误差可表示成如下形式:

$$g[u] = \int_0^T H(x(t), t) dt \quad (2)$$

我们的问题是: 由一系列观测数据 x_{obs} , 对模式初值 u 进行最优选择。我们从任一 u^k 出发, 沿着 $g[u]$ 在 u^k 处的负梯度方向, 即 $(-\nabla_u g)|_{u^k}$, 于是迭代序列为

$$u^{k+1} = u^k - \rho^k (\nabla_u g|_{u^k}) \quad (3)$$

设 u 扰动成 $u + ah$, 此时对应的解为 $\tilde{x}$, 令

$$\hat{x} = \lim_{a \rightarrow 0} \frac{\tilde{x} - x}{a}$$

则 $\hat{x}$ 满足如下线性方程

$$\begin{cases} \frac{d\hat{x}}{dt} = F'(x)\hat{x} \\ \hat{x}|_{t=0} = h \end{cases} \quad (4)$$

其伴随方程为

$$\begin{cases} -\frac{\partial P}{\partial t} = F'^*(x)P + \nabla_x H(x(t)) \\ P|_{t=T} = 0 \end{cases} \quad (5)$$

于是, 理论分析得到 $\nabla_u g = P(0)$, 从而求 $\nabla_u g$ 只需对(5)式反向积分到 0, 此时 $P(0)$ 即为 $\nabla_u g$ 。

下面对四维同化技术作几点说明。

说明一: (5)式为(4)式的伴随方程, $P|_{t=T} = 0$ 为(4)式的伴随边界条件。对(5)式反向积分, 不是不适定问题, 而是一个正问题。因为只须令 $\tau = -t$, 此时(5)式转化成

$$\begin{cases} \frac{\partial P}{\partial \tau} = F^*(x)P + \nabla_x H(x(\tau)) \\ P|_{\tau=T} = 0 \end{cases} \quad (6)$$

而 $\nabla_u J = P(0)$, 因为方程(6)的结构与(4)式的结构是相同的。故(4)式存在唯一解, 表示(6)式亦存在唯一解。于是通过对(6)式从 $-T$ 到 0 的正向积分可得到 $P(0)$ 的值。

说明二: 用最速下降法来求 u^* 时, 采用迭代格式(3)式, 如何正确选择步长 ρ^* 是一个关键的问题, 有关内容可参看 § 6.9 节和文献[35]。

说明三: 一般大气数值预报模式中含有两类方程, 一类预报方程, 另一类为诊断方程, 状态变量分成两部分 $x = (x_p, x_d)$, x_p 为预报变量, x_d 为诊断变量, 于是方程为

$$\begin{cases} \frac{dx}{dt} = F(x_p, x_d), G(x_p, x_d) = 0 \\ x_p|_{t=0} = u \end{cases}$$

问题转化为求 u^* , 使 $g[u] = \frac{1}{2} \int_0^T \|C \cdot x - x_{\text{obs}}\|^2 dt = \min$, 其中 x_{obs} 为观测资料, C 为线性算子。

另一类模式属于有限区域模式, 其模式优化问题是如何提供离散边界条件, 设有限区域模式在空间是离散的, 其模式如下:

$$\begin{cases} \frac{dx}{dt} = F(x) + B \cdot V \\ x|_{t=0} = u \end{cases} \quad (8)$$

其中 B 是时间的函数, 它定义在边界 $\partial\Omega$ 上, 仍为线性算子, 它把边界 $\partial\Omega$ 上的值与区域 Ω 内的值建立了联系, 寻找 u, v 使 $J[u, v] = \frac{1}{2} \int_0^T \|x - x_{\text{obs}}\|^2 dt = \min$ 。

其它还有离散模式, 这里不在赘述, 有下兴趣的读者可参阅有关的文献。

说明四: 四维同化中数据预处理问题。数据预处理技术有重大的作用, 它往往是数据处理成败的关键, 资料预处理的目的, 是要剔除错误资料, 对资料分级使用。

说明五: Kalman 滤波与四维同化技术。我们用 W_k^T 表示 k 时刻的真值, 一般是随机矢量, 假定用 $W_k^T = \Psi W_{k-1}^T + b_{k-1}^T$ 来表示前后两个时刻的真值用 Ψ 联系起来, 在一般 Kalman 滤波中记为 $\Psi_{k,k-1}$, Ψ 称为预报算子, b_{k-1}^T 表示预报误差, 一般假定为白噪声, 相当于 Kalman 滤波的信号方程。用 W_k^o 表示 k 时刻的观测值, 用方程 $W_k^o = H W_{k-1}^T + b_{k-1}^o$ 来联系起来, 在考虑线性情况时, H_k 为矩阵, 事实上 H_k 就是最优内插得到的内插算子, b_k^o 为观测误差, 一般假定为白噪声。上述方程称为 Kalman 滤波的观测方程, 自然假定 $\{b_k^o\}$ 与 $\{b_k^T\}$ 相互独立。Kalman 滤波不是去取代原有内插公式, 而是在原有数值预报, 最优内插的基础上, 进一步合理利用信息, 改善预报, 利用 Kalman 滤波模型, 就可充分合理利用有效信息, 与以往最优内插相比较, 它更有一个时间发展信息的利用, 把 Kalman

滤波应用于四维同化,还有一系列问题需要解决,如 Ψ, H_i 如何给?公式初值如何给等问题。有兴趣的同志可参考文献[36]。

说明六:有限域(即带有边界)上的四维同化技术。这里有几个问题:边界上的值如何给?边界条件给法必须符合物理背景,它亦直接影响到伴随方程的形式;另外伴随方程必须给出相应的伴随边界条件,这些条件的给出是由边界性质及方程性质所决定的。对于有限域上的问题,不仅对初值,而且对边值及模式中参数(包括函数)进行同化,这是一个相当困难和复杂的问题,但是对预报来说又是一个必须要解决的问题。

4. 变分与反问题的关系

在力学、地球物理、医学、工程技术、卫星遥感和气候预测等与国计民生和社会发展密切相关的数学物理中反问题在近几年已迅速发展起来,这类问题的显著特点是:不适定性和非线性性。特别是不适定性,使得通常的数值方法失败,故对应用带来相当大的麻烦。

对于一个系统,通常有三部分组成:输入信号 x (因),模型 K (关系),输出信号 g (果),于是, x, K 和 g 之间关系,可以用如下数学模型来表示:

$$Kx = g \quad (9)$$

其中 $K: H_1 \rightarrow H_2, H_i$ 为 Hilbert 空间。所谓正问题(direct problem)指:由因求果;所谓反问题(inverse problem)一般来说指:由果求因。这类问题往往是不适定的(ill-posed)。不适定的原因有三点: K 不满,即 $g \notin R(K)$,故(9)式问题解可能不存在; K 不单射,表示(9)式问题解不唯一; K^{-1} 存在但不连续,则(9)式的解 $x = K^{-1}g$ 不连续依赖于 g 。

对不适定问题而言,我们把精力放在解的稳定性问题上。设法构造稳定近似解来逼近“真解”,有两类问题需研究:

第一类问题: 设 $g \in R(K)$,此时(9)式可解,且设解亦唯一,于是解可表示成 $x = K^{-1}g$ 。由于 g 是测量得到,故有一定误差,记 g^δ 为 g 带有误差的观测值 $\|g - g^\delta\| \leq \delta$,于是我们实际上解 $Kx = g^\delta$ 的方程。若 $g^\delta \in R(K)$,此时用 $K^{-1}g^\delta$ 作为近似解,由于 K^{-1} 不连续,当 $\|g - g^\delta\| \leq \delta \rightarrow 0$ 时,不能保证 $\|K^{-1}g - K^{-1}g^\delta\| \rightarrow 0 (\delta \rightarrow 0)$,故 $K^{-1}g^\delta$ 不能担当稳定近似解的重任,必须设法构造一个近似解 x^δ ,使 $\|K^{-1}g - x^\delta\| \rightarrow 0 (\delta \rightarrow 0)$ 。

第二类问题: $g \notin R(K)$,此时解根本不存在或者我们根本无法判定 g 是否属于 $R(K)$,也不知道解是否唯一,基此原因,我们设法扩充解的概念,寻找(9)式的广义解 x^+ ,要求:如果(9)式的解存在的话, x^+ 就是解; x^+ 是存在唯一。于是把问题绕道走,避开了讨论问题解的存在和唯一性,只需寻找稳定近似解 $x^\delta \rightarrow x^+ (\delta \rightarrow 0)$ 。 x^+ 称为方程(9)的极小模最小二乘解,记为 $x^+ = K^+g$, K^+ 称为 K 的 Moore-Penrose 广义逆。广义解问题解决了,又面临新的问题。由于 K^+ 一般为无界算子,从而 K^+g^δ 不会收敛于 $K^+g (\delta \rightarrow 0)$,亦即用 K^+g^δ 作为稳定近似解是不合适的,于是,设法构造稳定近似解 x_α^δ 来逼近。

在不适应问题解法中,目前应用最广,理论上最完善的方法是线性不适定的正则化方法,这方法的一般理论由前苏联学者 Tikhonov 等人创立^[37],即引进 Tikhonov 泛函

$$M_{\alpha}[x, g^{\delta}] = \|Kx - g^{\delta}\|^2 + \alpha \|x\|^2 = \min$$

解得

$$x_{\alpha}^{\delta} = (\alpha I + K^* K)^{-1} K^* g^{\delta}$$

其中 α 称为正则化参数。

这里需要研究的问题:

①正则化中, α 的选择至关重要,它关系到近似解 x_{α}^{δ} 是否逼近于 $K^+ g (\delta \rightarrow 0)$ 。目前通常采用 Morozov 偏差原则(具体见 § 6.5 和 § 6.6),即 α 满足

$$\|Kx_{\alpha}^{\delta} - g^{\delta}\| = \delta$$

②提高收敛精度的问题,从而提出了一般的正则化原则。

③非线性不适定问题的研究。

④数值方法实施的问题。

有关问题这里不再一一介绍,有兴趣的读者可参阅有关文献[38~39]。关于反问题在卫星遥感上的应用可参阅文献^[40~42]。

5. 从近代数学的发展看变分问题

下面从几个阶段来看数学的发展与变分问题的关系:

第一阶段 从函数极值过渡到泛函极值

1696 年, J. Bernoulli 向各国数学家提出了挑战性的问题即著名的最速降线问题:求从 A 点到不是在它垂直下方的另一点 B(见图 1)的一条曲线 $y = y(x)$, 使得沿曲线在不计摩擦的情况下滑到 B(a, b) 点所花的时间最短。

由能量守恒可知

$$mgy = \frac{1}{2}mv^2 = \frac{1}{2}m\left(\frac{ds}{dt}\right)^2$$

其中 $ds = \sqrt{1 + y'(x)^2}dx$, 于是从 A 点到达 B 点所化时间为

$$T[y(x)] = \int_0^a \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2gy(x)}} dx$$

式中 $y(0) = 0, y(a) = b$ 。

要求 $y = y(x)$ 满足 $y(0) = 0, y(a) = b$, 且使 $T = \min!$, 这个问题与函数极值问题是两个完全不同类型的问题。函数极值问题: $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^1$, 求 $\Omega \subset \mathbb{R}^n$ 中一个点, 使 $f = \min!$, 而 Bernoulli 问题却完全不同, 记 $M = \{\text{连线 A、B 两点光滑曲线}\}$, 求极值即求 M 中一条曲线, 使 $T = \min!$, 这两个问题有着本质上差别, 后者人们称为泛函的极值。“泛函”顾名思义, 比函数更普遍、更广泛。泛函极值的产生, 引起人们极大的兴趣, 如 Newton, Leibniz, L'Hospital 等人都对此作过研究, 只有 Jacobi 所得结果与众不同,

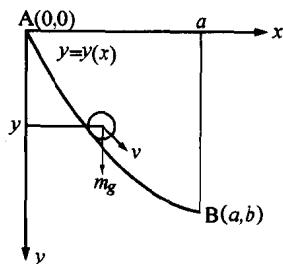


图 1

在他那“很不优美”的解答中,看到了其他人所没有看到的事实——从无穷多条曲线中选出一条满足极值条件的曲线,这个问题是**新的问题**。

第二阶段 泛函的极值与 Euler 方程的建立

Euler 以特有的与众不同的眼光研究了泛函的极值问题。他把泛函的极值与一元函数求极值巧妙有机地联系起来,为求泛函极值作出了极有价值的一步,即建立了泛函极值问题与微分方程边值问题之间的联系,此微分方程后来称为 Euler 方程。由于求解微分方程的边值问题比求泛函极值更为困难,于是人们又反过来,设法把某类微分方程边值问题转化为某泛函的极值来处理,这为解决数学物理中许多问题提供了一种极有效的方法,即建立了微分方程边值问题与泛函极值之间的等价关系,这就是**变分原理**。

第三阶段 求泛函极值到微分方程边值问题广义解的引进

下面从一个简单例子来说明。考虑泛函

$$J[w] = \int_{\Omega} \left[\frac{1}{2} |\nabla w|^2 - fw \right] dx \quad (10)$$

设 $w \in W \triangleq C^1(\Omega) = \{w | w \text{ 在 } \Omega \text{ 上一阶连续可微}\}$, $\Omega \subset \mathbb{R}^n$ 为开集, $f: \Omega \rightarrow \mathbb{R}^1$, 于是 J 定义了 $W \rightarrow \mathbb{R}^1$ 上一个泛函, 要求 $u \in W$, 使 $J[u] = \min$ 。利用 Euler 的思想, 若 u 使泛函 J 取得极小, 则对任何固定 $v \in W$, $t \in \mathbb{R}^1$, 作 $J[t] \triangleq J[u + tv]$ 在 $t = 0$ 点取得极值, 即 $J'(0) = 0$, 易得

$$\int_{\Omega} [\nabla u \cdot \nabla v - fv] dx = 0 \quad \forall v \in W \quad (11)$$

(11) 式即为 **Euler 方程的积分形式**

若 $u \in W$ 且 u 具有较好的光滑性(例如 u 是二阶连续可微), 则(11) 式可以分部积分一次, 如

$$\int_{\Omega} [-\Delta u - f]v dx + \int_{\partial\Omega} \frac{\partial u}{\partial n} \cdot v dx = 0 \quad \forall v \in W \quad (12)$$

第一种情形: 若 $u \in W$ 且 $\frac{\partial u}{\partial n} \Big|_{\partial\Omega} = 0$, 则 Euler 方程为

$$\begin{cases} -\Delta u = f & \text{在 } \Omega \text{ 中} \\ \frac{\partial u}{\partial n} \Big|_{\partial\Omega} = 0 \end{cases} \quad (13)$$

这时 Euler 方程所对应的边值问题为 **Neumann 问题**。

第二种情形: 若 $u \in W = C_0^1(\Omega)$ (粗糙地说, $v \in C^1(\Omega)$ 且 v 在 $\partial\Omega$ 附近为零), 则 Euler 方程为

$$\begin{cases} -\Delta u = f & \text{在 } \Omega \text{ 中} \\ u|_{\partial\Omega} = 0 \end{cases} \quad (14)$$

这时 Euler 方程所对应的边值问题为 **Dirichlet 问题**。

以上分析,我们得到二个重要启示:

启示一:变分问题在不同空间 W 中的解对应于不同微分方程的边值问题。

启示二:从 $J'(0) = 0$ 到导出 Euler 方程,我们作了一次人为的分部积分。本来 u 只需一阶可微,分部积分后对 u 需二阶可微,我们不能容忍人为的做法,所以必须降低 u 的可微性,即扩大定义解的函数类,“这不仅是为了数学本身的优美,而且也是为了使数学能更好地符合客观实际的需要”^[43]。我们把扩大了解的函数类以后的解称为广义解。而把原来二阶可微的解称为古典解。广义解的引进必须符合下面基本原则:古典解必为广义解;广义解是唯一的,且按某种度量连续依赖于初始定解条件。广义解的引进,使得古典变分产生一大飞跃,并且蕴育着近代数学的思路。

对上面具体问题而言,当 $W = C^1(\Omega)$ 时,称(11)式为(13)式的 Neumann 问题的广义解,若取 $W = C_0^1(\Omega)$,称(11)式为(14)式的 Dirichlet 问题的广义解。

第四阶段 泛函极值的可达性及 Sobolev 空间的引进

针对上面泛函在空间 W 中求极值,会产生新的问题,即在 W 中能否找到合适的 u , 使 $J[u] = \inf_{w \in W} J[w]$ 。这个问题并非是一个简单的问题,它意味着泛函极值是否能达到

的问题。

19 世纪中叶,数学中一个重大问题为 Dirichlet 问题。Riemann 曾提出了著名的 Dirichlet 原理,认为只要泛函有下界,就有下确界,此时变分问题就有解。Riemann 的观点只说对了一半。1870 年,Weierstrass 对 Riemann 论据提出了本质的批评,泛函有下界,则必有下确界,但下确界却不一定可以达到。一直到 Riemann 去世,尚未找到问题的答案,直到 Hilbert 才解决了此问题(1899,1901 的两篇论文)。在 Hilbert 看来,这个问题在数学中地位如此之重要,所以在 1900 年国际数学家大会上提出了著名的 23 个数学问题中,竟然有 3 个问题(第 19,20 和 23 个问题)与此直接有关,他提醒人们对这个问题以足够的重视,并第一个完整地证明了 Dirichlet 原理,从此开辟了变分的新纪元,后来数学物理的发展,完全证实了 Hilbert 的英明预见,直到四十年代发展了广义函数理论以后,Sobolev 空间成为近代数学与物理的重要工具(详见 § 5.5 和 § 5.6)。

第五阶段 近代变分的产生

近几十年来,变分原理有了重大的发展,形成了近代变分方法(又称为大范围变分)。近代变分方法(主要指临界点理论中极大、极小原理等),集线性与非线性泛函分析,偏微分方程及拓扑学等各学科知识,在拟线性与非线性问题中找到了广泛的应用。众所周知,微分方程边值问题的广义解与泛函极值相对应;在线性情况下,其广义解与泛函极小相对应,但在非线性情况下,其相应的泛函可能既没有上界,又没有下界。例如非线性问题

$$\begin{cases} -\Delta u = u^2 & \text{在 } \Omega \text{ 中} \\ u|_{\partial\Omega} = 0 \end{cases} \quad (15)$$

此问题化为泛函在空间 $W_0^{1,2}(\Omega)$ (粗糙地说, $u \in W_0^{1,2}(\Omega)$ 指 $u, \frac{\partial u}{\partial x_i} \in L^2(\Omega), u|_{\partial\Omega} = 0$ 的函数空间, 具体的意义见第五章 § 5.6) 中极值问题

$$J[u] = \int_{\Omega} \left[\frac{1}{2} |\nabla u|^2 - \frac{1}{3} u^3 \right] dx \quad (16)$$

但很可惜, $J[u]$ 既无上界, 又无下界, 原因是: 若取 $u \in W_0^{1,2}(\Omega), u \neq 0, u \geq 0$, 于是对任意 $t \in \mathbf{R}^1$, 均有 $tu \in W_0^{1,2}(\Omega)$, 此时

$$J[tu] = \frac{t^2}{2} \int_{\Omega} |\nabla u|^2 dx - \frac{t^3}{3} \int_{\Omega} u^3 dx \rightarrow \pm \infty \quad (t \rightarrow \pm \infty) \quad (17)$$

于是 1973 年, Ambrosetti 和 Rabinowitz^[44] 提出了“山路引理”, 它形象地说明: 从盆地中心出发, 到盆地外部, 必有一条道路从山脉的最低点越过, 这个最低点就是一个临界点。山路引理及各种山路定理的建立, 特别是它们在非线性问题中取得一系列新的结果, 使得临界点理论在几十年中取得重大进展, 有兴趣的读者可参阅专著^[45~47]。有关这方面的进一步发展, 例如临界点理论, 变分不等方程^[48,49]等可参阅有关方面专著。因为上述这类问题涉及到更多的数学知识, 故本书不再涉及。

二、本书内容安排

本书共分七章:

第一章介绍数学基本知识, 包括常微分方程的解法, 稳定性及定性理论, 偏微分方程的特征理论, 积分方程的理论及解法, 摄动的概念及奇异性摄动的几种方法介绍, 这些内容为后面学习变分方法是必不可少的。

第二章介绍数学物理中的变分问题, 共分六节, 主要介绍 $\mathbf{R}^n$ 中极值问题; 泛函的极值与变分, 泛函的条件极值与 Lagrange 乘子法; 力学中的变分原理, 指出分析力学中虚功原理是力学中变分原理的微分形式, 而 Hamilton 原理是力学中变分原理的积分形式; 介绍可变区域上泛函的变分与 Noether 定理, 研究泛函在坐标与函数变换下的不变性条件; 介绍流体力学中变分原理, 讨论变分原理及广义变分原理在流体力学、旋转流体中的应用; 最后介绍平均变分原理在流体力学及大气运动中的应用。

第三章介绍 Hamilton 力学与一阶非线性偏微分方程之间的关系, 共分八节。内容包括: 从几何观点讨论一阶拟线性及非线性初值问题求解的一般理论, 研究全积分及求全积分的方法, 对 Hamilton-Jacobi 方程来说, 如能得到全积分, 则对求解 Hamilton-Jacobi 方程是十分有利的; 介绍 Legendre 变换在力学上的应用, 利用 Legendre 变换可以把 Euler 方程与 Lagrange 函数转化为 Hamilton 正则方程及 Hamilton 函数, 在 Lagrange 函数中, 它只具有运动学意义, 而没有动力学的意义, 而引进的 Hamilton 函数, 它体现了动力学的意义, 而 Hamilton 正则方程形式简单且对称, 故它更具有物理特征; 在此基础上, 我们引进 Hamilton 主函数 S , 它作为判定系统真实运动准则的一个物理