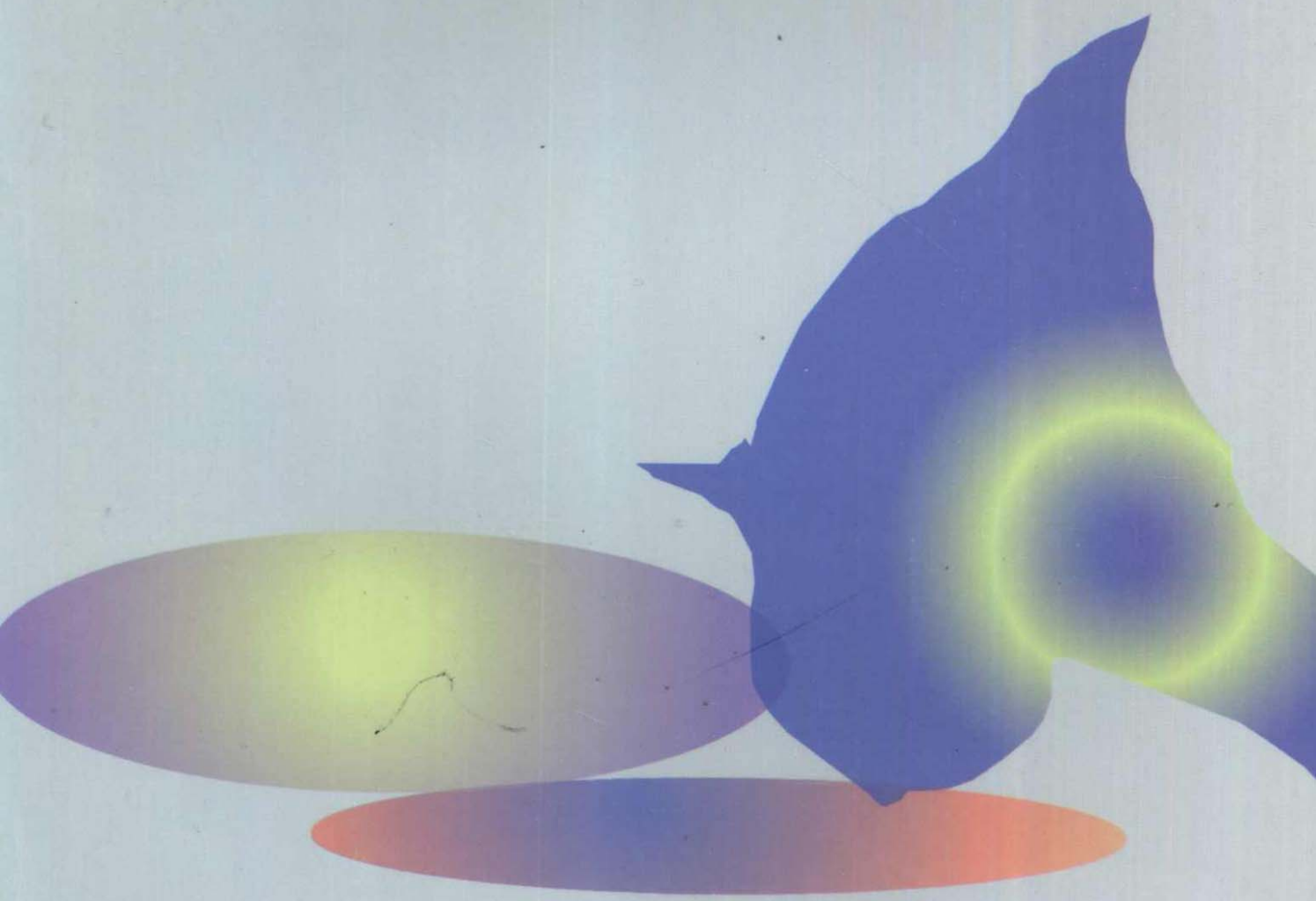


误差与数据处理 理论

周开学 李书光 编著



石油大学出版社

0241.1-43

274

误差与数据处理理论

周开学 李书光 编著



A1040073

石油大学出版社

图书在版编目(CIP)数据

误差与数据处理理论/周开学编著. —东营:石油大学出版社,2002.5

ISBN 7-5636-1647-0

I. 误… I. 周… III. ①物理-实验-误差-高等学校-教材 ②物理-实验-数据处理-高等学校-教材
IV. O.4-33

中国版本图书馆 CIP 数据核字(2002)第 027888 号

误差与数据处理理论

周开学 李书光 编著

出版者:石油大学出版社(山东东营 邮编 257061)

网 址: <http://sunctr.hdpu.edu.cn/~upcpres>

电子信箱: upcpres@mail.hdpu.edu.cn

印刷者:昌邑恒信印刷有限责任公司

发 行 者:石油大学出版社(电话 0546—8392563)

开 本:850×1168 1/32 印张:8.875 字数:230千字

版 次:2002年7月第1版第1次印刷

印 数:1—1500册

定 价:12.50元

序

误差与数据处理课是理工科各专业大学生的一门基础性课程。这门课程的开设,对学生学好各种实验课程、顺利阅读科技书籍和文献、做好毕业论文以及毕业后开展研究工作和科技交流都有非常重要的意义。近年来,选修该课程的学生越来越多,编写一本合适的教材就显得非常重要。周开学、李书光同志就是应这一需要编写了这本书。该书的出版对高校的教材建设是一个贡献。

这本教材是根据理工科大学的教学实际情况,本着深入浅出、便于学生学习的原则编写的。

该书的特点是:一、引入了适用于国际社会的关于测量结果的不确定度表示方法。二、编入的内容适合学时较短的课程教学。三、给出了较多的例题供学生参考。四、避开了过繁的推导而注重概念的讲授。五、设置了必要的阶梯,便于低年级学生学习。

综上所述,我认为该书是一本适用于除计量专业外的理工科各专业的好教材,相信会受到广大读者的欢迎。

孙在春

2002年3月25日

前 言

现代科学与技术工作离不开实验,而实验又离不开对具体量值的测量,并且测量总是存在误差的。人们总希望减少或消除测量中存在的误差来提高测量精度。误差理论正是研究误差的来源、性质、规律,减小误差的措施,评估测量结果精度的方法,如何合理设计实验去减小甚至消除误差,以及提高实验或测量精度的理论。它还能告诉我们在存在误差的情况下,如何在测量完毕之后估算待测量的大小和精度;指导我们采取正确的措施在测量过程中减小或消除某些误差;它还能指导我们为了减小误差的影响去正确地选择测量方法,选用合适的仪器,合理控制实验条件,采用正确的方法处理所得到的测量数据。

数据处理泛指对数据,特别是实验观测数据进行的一切加工及运算,包括由测量数据求待测量值,估算测量精度,列表或作图表示数据之间的对应关系与数据变化规律,以及通过所得到的数据求取相关未知量、求取相关量之间的未知关系式等等。该理论能帮助我们充分地利用测量得到的有限数据获得更多有用信息。

由以上可见,对于从事实验技术工作、计量工作的人而言,误差与数据处理理论是一项重要的基础理论。

为便于统一标准加强交流,自 1980 以来,国际计量大会多次推荐对测量结果采用不确定度表示其精度,我国国家计量部门积极响应,并且逐步在科技工作中推广国际上关于不确定度方面的建议。在当前对外开放逐步扩大的形势下,用不确定度对测量结果精度进行评定,已成为国际间科技交流及经贸往来的大趋势。

本书中,第 1 章介绍测量误差及有效数字等基本概念。第 2 章对随机误差、系统误差分别进行详细讨论,同时讨论了等精度、非

等精度情况下误差的特性及处理方法,包括减小或消除某些误差的方法。第3章讨论误差的合成,分别讨论直接测量中各种误差分量的合成和间接测量中各直接测量误差的合成(误差传递)。作为本章理论的应用,介绍了在设计实验时如何选用仪器、控制环境条件,比较论证测量方法优劣等等。第4章是不确定度知识,介绍了最新与国际接轨的评定测量结果精度的方法及报告方法。第5章介绍最小二乘法原理及用最小二乘法处理数据的方法,同时介绍了常用的列表、作图方法。第6章是回归分析知识,讨论如何由实验数据寻求相关量间的未知关系式,并对所得关系的可信度进行检验。第7章简单介绍了随机过程理论,为处理动态测量的数据打下一定的基础。

学习误差与数据处理理论,要用到数理统计知识和实验技术知识。为便于未学过上述课程的读者学习,对有关的实验知识做了稍为详细的叙述,在附录中给出了有关数理统计的基本概念和基础知识。书中编入了较多的例题帮助读者理解理论知识和方法。

本书是作为工科高校教材编写的,适合30至60学时课程教学使用。亦可供教师与从事相关工作的科技工作者参考。

书中第1~4章由李书光执笔,第5~7章由周开学执笔,相互之间也参与了一些对方的编写工作。编写该书过程中得到了多方支持,在此表示诚挚的谢意。书中难免存在疏漏或谬误,恳请指教。

编者

2002年3月15日

目 录

第 1 章 基本概念	(1)
1.1 研究误差的意义	(1)
1.2 测量的基本概念	(2)
1.3 误差的基本概念	(5)
1.4 精度的概念	(11)
1.5 有效数字	(13)
习题一	(17)
第 2 章 误差的性质及处理	(19)
2.1 随机误差	(19)
2.2 系统误差	(52)
2.3 粗大误差	(68)
2.4 直接测量的数据处理步骤	(79)
习题二	(83)
第 3 章 误差的合成	(86)
3.1 误差的相关性及相关系数	(87)
3.2 间接测量的误差	(92)
3.3 误差的合成	(106)
3.4 间接测量的数据处理步骤	(114)
习题三	(115)
第 4 章 测量不确定度	(117)
4.1 测量不确定度的基本概念	(118)
4.2 测量不确定度的来源及分类	(119)
4.3 测量不确定度的评定	(122)
4.4 不确定度的合成	(129)

4.5	不确定度的报告	(133)
4.6	测量不确定度应用举例	(135)
	习题四	(142)
第5章	数据处理常用方法	(144)
5.1	列表法	(144)
5.2	作图法	(151)
5.3	最小二乘法	(169)
	习题五	(194)
第6章	回归分析	(196)
6.1	回归分析基本知识	(196)
6.2	简易回归方法	(199)
6.3	最小二乘法一元线性回归	(207)
6.4	多元线性回归	(223)
	习题六	(230)
第7章	随机过程简介	(232)
7.1	随机过程及其特征	(232)
7.2	随机过程特征量的估计	(240)
	习题七	(252)
附录1	概率论与数据统计基本知识	(254)
附录2	正态分布的概率密度函数	(264)
附录3	附表	(268)
	参考文献	(275)

第 1 章 基本概念

1.1 研究误差的意义

人类对自然界所发生的量变现象的研究,常常要借助于各式各样的实验与测量来完成。然而,由于认识能力的不足和技术水平的限制,实验中测得的值和它的真值并不一致,这种差异在数值上表现为误差。随着科学水平的提高和人们的经验、技巧、专门知识的丰富,误差虽可以被控制得越来越小,却始终不能把它消除至零。测量结果都有误差,误差存在于一切科学实验和测量过程中,这已成为一条公理。

既然误差的存在具有必然性和普遍性,而这种特性又不可避免地影响实验测量结果,因此,要得到真实的、有价值的实验结果,就应该充分地认识误差,并设法减小或消除它,这也正是研究误差的意义所在。

误差理论及数据处理的主要任务是:

① 研究误差的性质,分析误差产生的原因,找出减小或消除误差的方法。

② 正确处理数据,合理评定所得结果的可靠程度,得到更接近于真实值的数据。

③ 正确组织实验,合理设计仪器或正确选择仪器和测量方法,力图在最经济的条件下,得到最理想的结果。

1.2 测量的基本概念

测量是科学实验的前提,误差与测量是紧密联系的。因此,在研究误差之前,有必要弄清测量的一些基本概念。

1.2.1 测量的定义

测量就是人们借助专门设备,以确定量值为目的所进行的操作。它是一个实验的比较过程,即通过实验把一个量(被测量)和另外一个量(标准量)相比较。关于测量,涉及下面两个概念。

1. 测量过程

执行测量所需的一系列操作。测量有时比较简单,可以用量器的单位量直接与被测量比较得到被测量值。有时不能直接进行比较,而必须通过仪器或某些辅助设备间接得到被测量值。因此,测量过程常常包括建立单位、设计工具、设计测量方法、研究分析测量结果、寻求消除或减小误差的方法等。

2. 测量结果

由测量所获得的被测量的值。一般情况下,测量结果由三部分组成,被测量值、单位及结果精度的评定。测量值实质上是被测量与其计量标准(单位)的比值;单位一般采用国际单位;而结果的评定采用不确定度,第4章对此有详细介绍。

1.2.2 测量的分类

按获取测量结果的方法,或测量条件及测量结果的不同,测量具有不同的分类法。本教材从三个不同的角度进行分类。

一、根据测量结果划分

1. 直接测量

把被测量与作为测量标准的量直接进行比较,或用预先按标准校对好的测量仪器对被测量进行测量,直接读到被测量数值大

小,这种测量为直接测量。用式(1-1)表示:

$$y = x \quad (1-1)$$

式中, y 是被测量的值, x 是其直接测量结果。

在工程测量中,用刻度尺测长度、钟表测时间、砝码与天平测质量、温度计测温度、电表(电压表和电流表)测电压和电流、频率计测频率等等,都属于直接测量。

2. 间接测量

先测量一个或多个直接测量值,然后利用已知的函数关系运算得到被测量,这种测量称为间接测量。间接测量可用式(1-2)的一般公式表示

$$y = f(x_1, x_2, \dots, x_m) \quad (1-2)$$

式中, y 是被测量, x_1, x_2, \dots, x_m 是直接测量的量值。

实际测量中,如果用直接测量的方法不能得到被测量,或能够测量但测量过程比较复杂,常常采用间接测量。例如体积、密度等量的测量往往用间接测量。

3. 组合测量

通过测量所有被测量的各种组合,列方程求解被测量的测量方法,称为组合测量。组合测量时,可以使各被测量以不同的组合形式出现,然后根据所得的结果,建立一组方程,通常采用最小二乘法进行处理(详见第5章)。例如测量几个电阻或电容,除了可以单独测量每个电阻或电容外,还可能通过测量它们的组合值,如串联值、并联值等来计算每个电阻或电容值。

二、根据测量条件划分

1. 等精度测量

在相同的测量精度条件下(同一测量水平的观测者、同一精度的测量仪器、同样环境条件、同样实验方法等)对同一待测量所做的重复性测量,称为等精度测量。

等精度测量所获得的数据,不能判定哪一个数据更接近于真

实值,只能一视同仁,同等对待,即对所有数据的信赖程度是相同的。

实际测量中,等精度的测量条件很难满足。一般情况下只是近似地认为是等精度测量。

2. 非等精度测量

在不同的测量精度条件下(不同测量水平的观测者、不同精度的仪器、不同方法、不同环境等)对同一待测量所做的测量,称为非等精度测量。

非等精度测量获得的各个数据,由于测量条件不同,因此,不能一视同仁,应区别对待,即对各数据的信赖程度是不同的。

三、根据观测对象在测量过程中所处的状态划分

1. 静态测量

被测量在测量过程中是固定不变的,对这种被测量进行的测量称为静态测量。

静态测量不需考虑时间因素对测量的影响,把被测量或误差作为随机变量来确定。

2. 动态测量

被测量处于随时间不断变化的状态,对这种被测量进行的测量称为动态测量。

动态测量要考虑时间对测量的影响,即把测得值或误差作为随机过程来进行研究。

随着科学技术的发展,工程技术中越来越重视动态测量及其数据处理。因为动态测量数据中,包含着大量有关被测量、测量器具以及外界环境加入的干扰等方面的信息,正确分析和处理动态测量数据,能得到许多反映客观事物规律的有用信息,所以动态测量数据处理是误差理论与数据处理的重要组成部分。

1.3 误差的基本概念

1.3.1 误差的定义

所谓误差是指被测量的给出值与该量的真值之差。可用下式表示。

$$\text{误差} = \text{给出值} - \text{真值} \quad (1-3)$$

式中,被测量的“给出值”是指通过某种方法确定的该量的数值大小,包括测量值、仪表示值、标称值及计算近似值等。真值是指在某一测量条件下该量所具有的客观真实值的大小,在某一时刻是一个确定值。由于误差存在的必然性和普遍性,一般情况下被测量的真值是不可知的。只有在某些特定情况下,真值才是可知的,具体如下:

1. 理论真值

通过理论方法获得的真值。例如,平面三角形三个内角之和恒为 180° ;理想电容或电感构成的电路,电压与电流的相位差为 90° ;理论设计或理论公式表达值等等。

2. 计量学的约定真值

国际计量机构通过内部约定而确定的真值。国际计量大会决议以下 7 个 SI 基本单位量。

① 长度单位——米(m)

光在真空中 $1/299\,792\,458\text{ s}$ 的时间间隔内所走过的路程定义为 1 米(m)。

② 时间单位——秒(s)

1 秒(s)是铯—133 原子基态的两个超精细能级之间跃迁所对应的辐射的 $9\,192\,631\,770$ 个周期电磁波所持续的时间。

③ 电流强度单位——安[培](A)

真空中,恒定电流通过两根相距 1 m 长而细的平行导线,产

生电磁场。如果两导线之间产生的力,在每米长度的导线上为 2×10^{-7} 牛顿时,此时流过导线的电流为 1 安[培]。

④ 质量单位——千克(kg)

以保存在法国巴黎国际计量局的铂—铱合金圆柱体(国际千克原器)的质量为标准 1 千克。

⑤ 热力学单位——开[尔文](K)

水三相点热力学温度的 $1/273.16$ 为 1 开尔文。

⑥ 发光强度单位——坎[德拉](cd)

频率为 540×10^{12} Hz 的单色辐射光源,在给定方向上的辐射强度为 $1/683$ W 每球面度($\text{W} \cdot \text{sr}^{-1}$)定义为 1 坎[德拉]

⑦ 物质的量的单位——摩尔(mol)

摩尔是一系统的物质的量,该系统中所包含的基本单元数与 0.012 kg 碳-12 的原子数目相等时为 1 摩尔。

3. 标准器的相对真值

高级标准器的误差与低一级标准器或普通计量仪器的误差相比,为其 $1/5$ (或 $1/3 \sim 1/20$)时,可认为前者是后者的相对真值,这在实际测量工作中很有意义。

1.3.2 误差的来源

一个完整的测量系统,通常由 4 个部分组成:实验源(自然源、人造源,如光、电、力、热等)、实验体(待测样品)、观测系统(人眼、传感器、测量仪器)以及所处的环境(如温度、湿度、气压、振动、光照、电磁场等)。各组成部分咱实验过程中都会引起误差,因此,关于误差的来源,可以归纳为以下几个方面。

一、仪器、装置误差

1. 标准器误差

标准器是提供某个被测参数标准量值的器具。例如,标准量块、标准电池、标准电阻和铯原子钟等。这些标准器所提供的“约定

真值”仍然有误差。另外,在没有满足标准器要求的使用条件或标准器不够完善时,标准器也产生附加误差。

2. 仪器误差

测量仪器是指将被测量转化成可直接观测的指示值或等效信息的计量器具。无论仪器设计得多么完善,制造得多么精密,都会有不足之处,从而影响仪器的性能。元件老化、放大器零点漂移等原因也会产生误差。例如,阿贝比较仪、天平、电桥等比较仪器、温度计、秒表、检流计、压力表等指示仪器,都会产生误差。

3. 附件误差

为测量创造条件,使测量方便进行而使用的各种辅助配件,均属测量附件。例如电测中的开关、导线、电源、热源等都会引起误差。

二、环境误差

由于各种环境因素与要求的标准状态不一致而引起的测量装置和被测量本身的变化所造成的误差。例如,环境温度的改变,湿度、大气压力、重力加速度的变化,电磁场干扰、外力冲击与环境振动等引起的误差,都属于环境误差。

仪器仪表在出厂规定的正常工作条件下使用时具有的误差称为基本误差,超出此条件使用时所增加的误差称为附加误差,也属于环境误差。

三、方法误差

由于测量方法或计算方法不完善、不合理等原因引起的误差,称为方法误差。例如,电测中引线电阻上的压降、阻抗不匹配;需要瞬时取样测量,而实际取样间隔时间不为零;经验公式函数类型选择的近似性和公式中各系数确定的近似性等引起的误差,均属于方法误差。

四、人员误差

人员误差是指测量人员分辨力有限,感觉器官的生理变化,反

应速度和固有习惯等主观因素引起的误差。例如,记录某一信号时,有滞后或超前的趋势;对准某状态读数时,始终偏左或偏右,偏上或偏下;偶尔读错、抄错数据等产生的误差均属此类误差。

1.3.3 误差的分类

根据误差的性质和表现形式,可将误差分为系统误差、随机误差和粗大误差3类。

一、系统误差

在同一测量条件下,多次测量同一量时,误差的绝对值和符号保持恒定;或在条件改变时,按某一确定的规律变化的误差称为系统误差。例如,仪表及量具的零点不准、钢卷尺随温度变化等引起的误差都属系统误差。

从不同角度,系统误差可分成不同种类。按对误差掌握的程度,系统误差可分为已定系统误差和未定系统误差。已定系统误差是指误差绝对值和符号已经确定;而未定系统误差是指误差绝对值和符号未确定,但通常可以估计出误差范围。

按误差数值的出现规律,系统误差可分为不变系统误差和变化系统误差。不变系统误差是指误差绝对值和符号为固定的系统误差;变化系统误差是指误差绝对值和符号为变化的系统误差,如线性、周期性及复杂规律变化等。

二、随机误差

同一测量条件下,多次测量同一量时,误差的绝对值时大时小,符号时正时负,以不可预知的方式变化,这种误差称为随机误差。天平变动性、实验条件的波动等引起的误差,均属于随机误差。

随机误差存在于任何测量条件下,并且不可修正。对其影响只能进行统计估计。

三、粗大误差

粗大误差又称疏失误差,它是由于工作人员疏忽、仪器失灵等

原因所造成的超出规定条件下预期的误差。含有粗大误差的测得值明显偏离被测量的真值,在处理测量数据时,应首先检验出含有粗大误差的数据并将它剔除。

应该指出,系统误差是由测量过程中某一个突出因素变化所引起的,随机误差是由测量过程中多种因素微小变化综合所致。二者虽然性质不同,但并不存在绝对的界限。系统误差中的变化系统误差,本身具有某种随机性,数值较小时,与随机误差的界限不十分明确。随着对误差性质认识的深化和测试技术的发展,人们可以将本来是随机误差的某些误差分离出来作为系统误差处理,或把某些系统误差进行随机化,作为随机误差来处理。

1.3.4 误差的表示方法

一、绝对误差

用绝对大小给出的误差定义为该量的绝对误差,例如

$$\text{绝对误差} = \text{给出值} - \text{真值} \quad (1-4)$$

绝对误差是有名数,即带有单位的数。可正可负,其数值大小与所取的单位有关,如 1 mA 可写成 0.001 A。

绝对误差反映被测量值偏离真值的大小及方向。

实际工作中,还经常使用修正值的概念。修正值的定义为

$$\text{修正值} = - \text{误差} = \text{真值} - \text{给出值} \quad (1-5)$$

变换可得式(1-6):

$$\text{真值} = \text{给出值} + \text{修正值} \quad (1-6)$$

这说明,含有误差的给出值加上修正值后就可消除误差的影响,非常符合人们的逻辑思维方式。

二、相对误差

绝对误差与被测量真值的比值称为相对误差,即

$$\text{相对误差} = \text{绝对误差} \div \text{真值} \quad (1-7)$$

相对误差是无名数,通常用“%”表示。