

计算机在高能物理实验 活动中的作用

戴 贵 亮

在今日的科学实验活动中,计算机是必不可少的工具;以探索微观世界秘密为目标的高能物理实验,更是如此。在粒子对撞机上进行的高能物理实验,事例率高,数据量大,运行实验所需的费用也高。为了提高实验效率,保证实验的质量,从数据采集到数据分析都需要有足够能力的计算机系统的支持。计算机在高能物理实验活动的每个方面都起着重要的作用。

一、高能物理实验数据获取与处理

计算机参与高能物理实验活动的一个重要方面,是实验的数据获取与处理。这里首先介绍一下数据获取与处理过程的概况。

谈到数据获取,首先涉及的问题是数据来源问题。一般而言,数据来源于测量系统;在高能物理实验中,测量系统是庞大的探测器装置和大量的电子学设备。

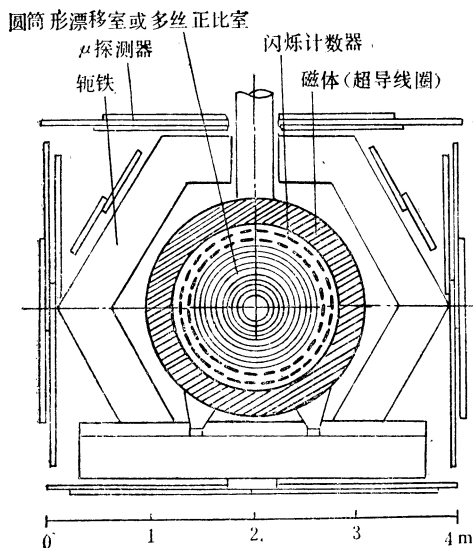


图1 探测器的示意图

图1给出了一个探测器的示意图(在与粒子束流

相垂直的面上的剖视)。在对撞机上进行高能物理实验所使用的通用型探测器,是一个体积很大的由大量多种类型探测单元构成的组合体,常用的探测单元有漂移室、簇射计数器、飞行时间探测器等;在这个组合体内,为了测量带电粒子动量和鉴别不同类型事例的需要,通常还安排有磁铁和吸收体。由于高能粒子束团对撞

产生的物理“事例”是围绕着对撞点四面八方分布的,所以,通常把探测器作成圆筒形;用大量的测量单元把对撞点严严实实地“包围”起来。

在高能物理研究中,感兴趣的物理量是粒子的质量、动量、电荷等等,这些量大多是不易直接测量的;探测器系统实际测量的是粒子的飞行轨迹、粒子的飞行时间、粒子穿过探测元件所留下的能量损失等等。每当在探测器包围的空间内有物理事例发生时,感受到事例的每个探测单元(也即粒子飞行经过的那些单元)

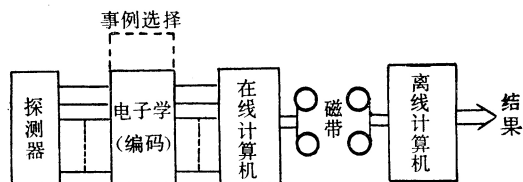


图2 数据获取与数据处理的基本过程

都将给出相应的电信号,这些信号,经过电子学电路的放大、编码,转换成数字代码,这就是实验系统最基本的数据来源。

图2用方框表示了高能物理实验中数据获取与数据处理的基本过程。在整个过程中,数据要经历“在线”与“离线”两种处理阶段;这两个阶段,均是由计算机完成的。由图中看到,在线计算机系统直接与探测器——电子学系统相连接,从而直接参加数据获取过程的活动。它的基本任务是从电子学电路读取已被编码的数据并将之记录到磁带上。由于实验数据的进一步分析所需的计算处理量极大,没有条件在实验进行过程中在在线计算机内随时完成,因而,记录在磁带上的“粗数据”的进一步分析,需在实验告一段落或至少是在一盘磁带已被记满之后,由在线计算机取下送到另外的具有更强处理能力的计算机系统上去进行。这种脱离了在线计算机在大型计算机上进行的数据分析处理过程叫做离线数据分析。

参照图 2, 可以将高能物理实验中数据经历的过程概括如下: 探测器直接与粒子束团对撞所发生的物理过程相接触, 在有物理事例发生时给出描述事例特征的电信号; 这些信号经过电子学电路进行编码形成数字量代码; 在线计算机负责读取已被编码的数据并将之记录到磁带上; 然后, 粗数据磁带被送入大型计算机作完全的数据分析, 经过大量的计算处理过程, 最终给出物理结果。

二、在线计算机系统

读取数据并将数据记录到磁带上, 是在线计算机系统的基本任务, 但这并不是它的全部任务。在整个实验系统中, 在线计算机实际是一个集中的控制和监督环节; 它要负责整个实验过程的管理, (如启动、停止、某些条件的设置与更动等), 负责对数据质量和探测器设备运行状况进行监督, 负责对磁场、气体、温度进行监测与控制等。此外, 在甚多的设计中, 在线计算机还承担“数据过滤”任务。下面对监督和数据过滤两项任务作一些说明。

1. 数据质量和设备运行状况的监督 高能物理实验的规模庞大, 支持实验所需的费用很高。为了确保实验数据的质量, 在数据获取的过程中, 需随时对数据质量进行监督, 对设备的工作状况进行监督。

设备的工作状况包括加速器束流条件和探测器系统的工作状况。对设备状况进行监督的主要办法是, 在数据获取过程中对数据作实时抽样和分项统计。抽样的频度和具体作统计的参量, 则根据实验要求和监督的目标的不同而异。譬如, 为了粗略监督束流的有无和探测器是否工作, 并不需很多统计数据, 只需以一

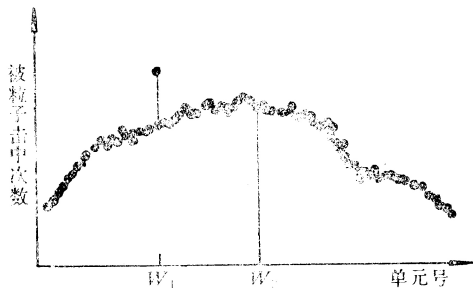


图 3 探测器工作单元被粒子击中次数的统计图

定频度扫视一下, 当发现异常时, 及时给出报警。而对于探测器各单元的效率, 则往往要对每个探测工作单元被粒子击中的次数作统计, 只有通过足够的累计数据, 才能查觉各单元效率的差异。计算机需安排在数据获取过程中同时对探测器各单元被粒子击中的次数进行统计, 并允许实验人员随时请求显示统计图。图 3 给出一个统计图的例子。横坐标是单元号, 纵坐标是累计的击中次数。在统计图上出现突出点, 说明相

应的单元发生了故障。例如, 单元 W_1 的统计数字过高, 说明该单元或相应电路有噪声; W_2 的统计数字为 0, 说明该单元或相应电路已不工作。

数据质量的监督是指对数据的物理质量进行监督, 这种监督帮助实验人员了解正在获取的数据是否为进一步感兴趣的数据。为了实现这种监督, 在线计算机系统需对获取的数据抽样, 并作更接近完全的数据分析, 其中包括重建事例的径迹和求出各种分布。经过分析所得到的结果可用统计图的方式或以单事例图象的方式在终端设备上显示出来。尤其是单事例显示, 是很直观的; 不只给出事例径迹的图象, 同时还给出主要探测器的外形。(图 4 给出了一个离线数据分析单事例显示的示例, 在线系统的单事例显示与此也很相似。)

2. 数据过滤 高能物理实验的一个特点是本底率高, 好的事例率相对较低; 尤其在对撞机上进行实验, 这个特点更加明显。所谓本底, 是指不感兴趣的那些事例: 如束团撞到对撞机管道壁或管道内的残余气体而产生的事例, 宇宙线通过探测器造成的事例等等。一般说来, 本底事例率可高达 10^6 次/秒或更高; 而真正感兴趣的好事例可能每秒不到 10 次或更低。而且, 在这些好事例中, 包含有多种类型的事例; 对于具体研究目标来说, 具体感兴趣类型的事例可能少到每秒出现 1 次甚而几小时才出现 1 次。也就是说, 本底与好事例的比率可能高达 10^6-10^8 以上。

高能物理实验就是要在高本底的大海中去寻找那些感兴趣的事例。所以, 原则上讲, 由探测器开始到得到物理结果为止的整个过程, 实际也是一步一步地对事例进行选择(或称判选)和对数据进行过滤的过程。选择的过程是逐级进行的, 在不同级次上使用不同的手段和不同的选择标准。例如, 在电子学设备中, 使用电

表 1

实验室	探测器名称	在线计算机系统	备注
SLAC	MARKII	一台 VAX-11/780	
	DELCO	一台 PDP-11/40 六台 LSI-11	四台用于监测
	MAC	一台 VAX-11/780 一台 LSI-11	LSI-11 用于监测
DESY	ARGUS	一台 VAX-11/780 一台 PDP-11/45	PDP-11/45 用于监测
	CELLO	一台 PDP-11/55 一台 PDP-11/45 九台微处理机	微处理机用于监测
	MARKJ	一台 PDP-11/55 二台带微处理机的设备	

子学逻辑电路完成判选。由于时间的限制和电路数量的限制,选择标准比较粗糙,在这一级上只期望滤掉经简单判别可以识别的坏事例;在在线计算机中,有条件使用比电子学更细一些的标准对事例进行检验,因而能进一步滤掉那些经过稍复杂一些的判断可以识别的坏事例;最精细的事例选择,需最后在离线数据分析中进行。由于离线使用的大型计算机(以下简称离线计算机)有更强的处理能力,而且,离线又不像在线那样受到时间的限制,所以可以使用更加精细的选择标准。

通常高能物理实验使用的在线计算机是具有快速处理能力的小型计算机。而且,在很多在线系统的设计中,使用两台或多台小型计算机构成多机工作方式,把各种任务分配到分别的计算机中去完成。表1列举了美国斯坦福直线加速器中心(SLAC)和西德电子同步加速器中心(DESY)对撞机上几个高能物理实验使用的在线计算机系统的构成概况。

在表1中并没反映出计算机外部设备的配置情况。实际上,外部设备的配置是很关键的,它不只直接关系到系统的性能,也对系统的成本有重要的影响。表2给出了一般在线计算机系统配置的外部设备的情况,其中包括各类外部设备的台数和用途的说明。

表 2

外部设备名称	台数	关于用途的说明
磁盘	1—2 台	用于存放系统的软件和记录数据
磁带机	2 台	用于记录数据
终端设备	1—2 台	用作实验人员的控制台
显示器与图形显示器	2—6 台	用于显示各种曲线,图表以及描绘探测器形状和事例特征的各种图形
图形输出设备	1—2 台	用于把显示的图形印刷出来,以便保存
打印机	1—2 台	打印数据
其它(如卡片机等)		

三、实验数据分析

从磁带的粗数据开始到最终的物理分析为止的整个处理过程统称为数据分析过程,这个过程是在中型或大型计算机上进行的。数据分析的基本任务是把粗数据转化为物理数据;处理的主要内容是径迹识别和径迹拟合。

径迹识别的过程是在粗数据中寻找径迹的过程,这个过程要求的处理量极大。在线计算机记录到磁带上的数据包含几方面的内容:与径迹空间坐标相关的粗数据、有关探测器部件相对几何位置的数据、磁场数据、以及探测器与电子学设备的标定数据等等,这些数据都是在径迹识别处理中需要使用的。在粗数据磁带上,描述径迹空间坐标的粗数据是离散的,这些数据并没有按每条径迹分组。如果形象地用图形表示,这

些数据就像一批彼此无连系的零零星星的坐标点(参看图4)。径迹识别的任务就是用数学的办法把这些离散点连接成“合理的”径迹。(也就是说,连接成的径迹应该是合乎物理规律的)。可以想像,在径迹条数较多,粗数据内夹杂有“噪声”的条件下,这种识别径迹的过程是十分麻烦的。这种处理往往不可避免地要经过各坐标数据反复组合反复试探的过程。在大型计算机中,重建一个事例通常也要占用半秒钟的处理时间,因而,要处理由在线计算机获取的大量事例数据,加给离线计算机系统的负担是很可观的。

图4示意地给出了粗数据离散点和经径迹识别后已连接成线段的若干条径迹的情况。该图也反映了在计算机图形显示器上能直观地显示探测器主要外形轮廓和显示事例径迹的情况。

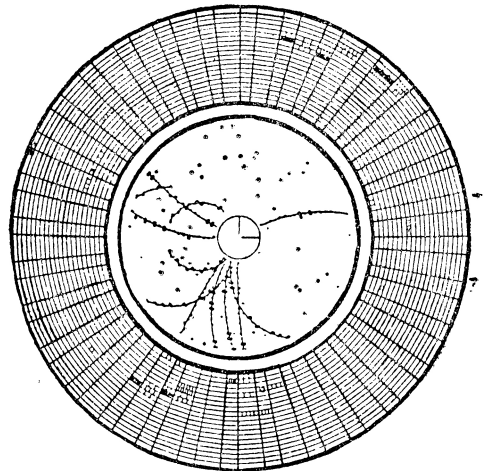


图4 把粗数据离散点连接成径迹

径迹拟合处理是在径迹识别之后进行的步骤。这个步骤的目标是针对每一条已被识别的径迹寻找出与该径迹数据最吻合的一条能反映物理规律的曲线,并确定出与这条曲线相应的一组物理参数。径迹拟合所需要的计算量也是极大的。

径迹识别与径迹拟合是数据分析过程中的两个基本步骤。在完成这两个步骤的基础上,可以进行粒子分类,事例选择,计算各种分布等处理,从而最终给出数据分析的结果。

四、物理实验的模拟

前面已介绍了由数据获取到数据分析的大体过程。如果概括地把这个过程抽象,可以认为,探测器与电子学设备是一套测量设备,它完成的作用是把发生于对撞点处的物理过程转换成以电信号形式存在的数字量;在线计算机与离线计算机是一套处理装置,它通过程序把上述的数字量处理成描述物理过程特点的数据;如粒子相互作用的截面,各种分布等等。这些数据再进一步经过理论工作者的加工,最后形成关于粒

子内部结构或内部相互作用的模型,或者说,形成对微观世界的某些认识。

从上述这些过程来看,除了在测量装置(即探测器与电子学)内进行的过程是“硬”过程之外,其余的过程(包括理论物理的计算在内)都是计算处理的过程。这些过程都是通过计算机的程序完成的。因而,我们容易想到:既然通过计算处理,能把“硬”过程给出的数字量处理成描述物理过程特点的数据,并且最后处理成描述理论模型的内部参数,为什么不能由理论模型内部参数开始,通过程序处理以相反的发展方向推算出反映物理过程的数据,甚而再向前推算出某些原始数字量呢?这无疑是可以的。这种以相反的发展方向处理的过程就是“模拟”。因为模拟的过程很象是用计算机的程序来取代真正的实验测量的过程,因而有人称为“计算机实验”。

在高能物理实验中,模拟是一项很重要的工作,是一项由开始着手准备实验就开始,而且只要还打算进行物理实验就不能停止进行的工作。由于实验的模拟可以在暂时“摆脱”硬过程的情况下作实验,因而,它既是一个检验实验方案可行性的手段,也是一个验证测量设备设计合理性的手段。最终,它更是预见实验结果的手段。实际上,这种方法是以前实验检验理论和以理论指导实验的一个枢纽。

高能物理使用的模拟方法是蒙特卡罗方法。这是一种数学统计的处理方法。使用这种方法可以针对不同的物理模型计算出相空间的分布,从而产生出模拟各种事例的数据。在探测器设计中,可以使用这种“蒙特卡罗事例”的数据通过计算来检验探测器方案的效果;在物理实验设计中,也可利用“蒙特卡罗事例”的数据来计算实验需要延续的时间,可能达到的精度等等。蒙特卡罗模拟对于避免实验的盲目性来说,是非常关键的。

五、小 结

以上诸节中,已重点介绍了在高能物理实验活动

表 3

应用的方面	计算机类型
实验在线系统	单独的小型计算机
探测器部件性能测试	单独的小型计算机或微型计算机
实验数据分析	大型计算机系统
实验模拟	
各种计算	
程序准备	
其它	

中需用到计算机的几个方面。但应该说明,计算机的应用方面并不止于此。例如,探测器的设计,探测器的性能测试,电子学电路的设计与性能测试等,都需要计算机的辅助。而且,计算机的程序设计与调试本身,也需在计算机上进行。概括起来,不同方面所需使用的计算机类型可参看表 3。

这里需再提一下的是大型计算机系统的问题。为了充分发挥系统的效能,为了提高系统的工作可靠性和改善使用者的使用环境,近代的大型计算机系统大多是由多台大型计算机互相连接构成网络,(可能不只限在一个地区,而是与全国性的大型计算机网络相连。)并且向外连接大量终端设备。这样,可以允许很多使用者同时共享整个计算机系统的资源(即计算机系统的全部潜力和外部设备)。允许他们在各自占用的终端上作数据分析、实验模拟等各种工作。国外各高能实验中心都设有相当规模的计算中心。计算中心设置的大型计算机系统除承担着表 3 中所列的与物理有关的处理任务外,还承担着与加速器设计有关的计算任务以及行政事务管理任务。为了说明高能实验中心大型计算机系统的规模,表 4 列举了美、欧几个实验室计算中心主要计算机构成的情况。

表 4

实验室	计算中心主要计算机	平均速度	备注
SLAC (美国)	IBM 3081 1台	1000 万次/秒	
	IBM 370/168 2台	500 万次/秒	
Fermi Lab. (美国)	CYBER 175 系统	1000 万次/秒	
DESY (西德)	IBM 370/168 2台	500 万次/秒	终端设备 40 台
	IBM 3033 1台	400 万次/秒	
CERN (日内瓦)	CDC 7600 1台	1500 万次/秒	终端设备 600 台
	CYBER 170-720 1台	140 万次/秒	
	CYBER 170-730 1台	200 万次/秒	
	IBM370/168 1台	250 万次/秒	
	IBM 3032 1台	250 万次/秒	

目前,我国已着手建造正-负电子对撞机,没有疑问,无论在对撞机设计方面还是在物理实验的准备方面(包括物理实验的模拟工作、探测器设计和性能测试工作等)都需要计算机的支持。我们现有的计算机设备条件是极差的。因而,建立自己起码的计算机系统是一件当务之急的事。与国外高能实验中心相比,当前我们的规模是比较小的。因而,没有必要建立像表 4 中列举的规模的计算机系统。但是,如果没有一个起码的计算机系统,会给对撞机的建造和物理实验的准备工作带来多方面的困难。