

· 科学论坛 ·

## 公共部门绩效评估:理论与实践<sup>\*</sup>

吴建南

(西安交通大学公共政策与管理学院, 西安 710049)

**[摘要]** 本文分析了公共部门绩效评估的基本内涵,并与绩效测量和绩效管理进行了比较,给出了包括绩效评估的若干要素及其相互关系的理论框架,分析了绩效评估的“绩效”与创新路径,并对科技计划绩效评估实践进行了简要分析。

**[关键词]** 绩效评估,绩效测量,绩效管理,公共部门,理论框架

随着公共部门的规模 and 影响在整个社会中所占的比重越来越大,对其绩效评估的关注在国内外均日益增强。在我国,国家领导人多次强调“抓紧研究建立科学的政府绩效评估体系和经济社会发展综合评估体系”,“加快行政管理体制改革,推行行政问责制度和政府绩效管理制度”等。<sup>1</sup> 2005年,财政部出台《中央部门预算支出绩效考评管理办法(试行)》,进一步推进了财政支出绩效评估的探索。目前,我国公共部门绩效评估实践已取得长足发展。地方政府主要采取自上而下的目标责任制考核和以政风行风评议和“万人评议”为主的公众参与绩效评估<sup>[1]</sup>,公共项目和财政支出绩效评估集中在科技、教育、农业等领域,如科研项目绩效评估、农业财政支出绩效评估等。

### 1 概念厘定与关系探讨

“人是评价的动物”,总在有意无意地进行价值判断。绩效,即某个对象干得好坏、行为表现乃至运行状态,包括感知绩效和客观绩效<sup>[2]</sup>。个人、团队、组织、项目、财政支出、系统等都有特定的绩效。绩效是对于干得好坏或行为表现的整体描述,是一个多维的概念。不同利益相关者在不同时期会关注绩效的不同方面。政府在不同时期有着不同的绩效侧重点,如在改革开放30年里,效率与公平如何平衡的问题始终为人们关注<sup>[3]</sup>。

目前常从四个维度(4E)来理解绩效,即经济性(Economy)、效率性(Efficiency)、效果性(Effectiveness)和公平性(Equity)<sup>[4]</sup>。经济性是成本标准,说明做某件事情花了多少钱;效率性是生产力标准,表示单位投入带来的直接产出;效果性即质量标准,用于衡量服务质量及是否达到了预期目的;公平性即公平标准,指机会、效果和努力在社会群体中的合理分配程度。此外,随着人们对公共产品和服务的要求越来越高,回应性(Responsiveness)、透明、问责(Accountability)等也成为绩效的诸多维度之一。

绩效评估、绩效测量、绩效管理是文献中经常出现的术语,但三者含义不同<sup>[2]</sup>。绩效评估是对特定对象行为表现的价值判断,考试、比赛、评选达标、表彰奖励等都是绩效评估的表现形式。绩效评估可以是显在的,也可以是潜在的。显在的绩效评估是正式规则下的行为,如组织对个人绩效的考核;潜在的绩效评估则更多地表现为领导者的基本能力,表现在其言行中,反映在奖惩上。潜在的绩效评估看不见,犹如“潜规则”,但其影响可能更大。

绩效测量是指对某一段时间系统运行状态的测定,包括指标设定、数据收集等一系列过程。绩效测量是对特定系统绩效的客观描述,是绩效信息的收集过程,理想情况下不包含主观的价值判断和感情色彩,一定程度是对绩效的还原和再现,是对现实世界特定对象行为表现的“模拟”过程。绩效测量可以

本文于2009年3月4日收到。

<sup>\*</sup> 国家自然科学基金项目和教育部新世纪优秀人才资助项目。

<sup>1</sup> 分别出自:温家宝总理在十届全国人大三次会议上所作的《2005年政府工作报告》(2006年3月5日)和在十一届全国人大一次会议上所作的《2007年政府工作报告》(2008年3月6日)。

分为日常测量和非日常测量。日常测量是边工作边提供绩效信息,而非日常测量是为了达到特定目的的事后专门测量。

绩效是测不准的。物理学中的“测不准原理”由海森堡于 1927 年首先提出,是量子力学的基本原理。该原理表明一个微观粒子的某些物理量不可能同时具有确定的数值,因而是测不准的。实际测量中,误差不可避免;在管理中,绩效测量的对象往往是具有高度复杂特征的组织和个人,在给定条件下,更难以完全准确地测量绩效。绩效评估的主体是人,其价值判断是主观行为,因此,绩效不仅难以准确测量,更无法客观评定——尽管我们总在努力采取多种方式降低测量误差、提高测量精度,乃至至于提升评价的信度与效度。为此,应谨慎使用绩效信息,慎做具体排名,采用多种途径测量和评估绩效,使结果相互印证。

绩效管理包括绩效计划、执行、评价、监督、报告、反馈、改进等一系列过程,目的在于监测、诊断和持续改进绩效<sup>[5]</sup>。绩效管理不只关注系统运行的结

果,而且关注整个系统从输入、过程到输出、结果和影响的全部内容。如果说管理是通过资源配置实现预定目标,而绩效管理则通过对整个系统的调节、配置,全面考虑经济性、效率性、效果性以及公平性从而实现短期与长期、局部与整体的科学发展目标。所以说,绩效管理在很大程度上就是管理,而管理在很大程度上就是绩效管理。

尽管内涵不同,但三者密切关联(如图 1 所示)。绩效评估的前提是绩效测量,而绩效管理的前提之一是绩效评估。绩效测量是对特定对象系统运行状况的模拟、还原和再现,是绩效评估的基础。绩效评估往往反映为一种权力,可以形成评估对象的动力和压力,是一种相对有效的管理控制机制。因此,绩效测量回答“是多少”的问题,主要是指标及权重设置问题,是没有掺杂感情色彩的客观描述。绩效评估回答“多少是好”的问题,主要涉及标准设定问题,是客观测量基础上的主观价值判断,有好坏优劣之分。绩效管理回答“如何更好”的问题,关键是如何诊断和改进的问题,它要求一种整体性思维。

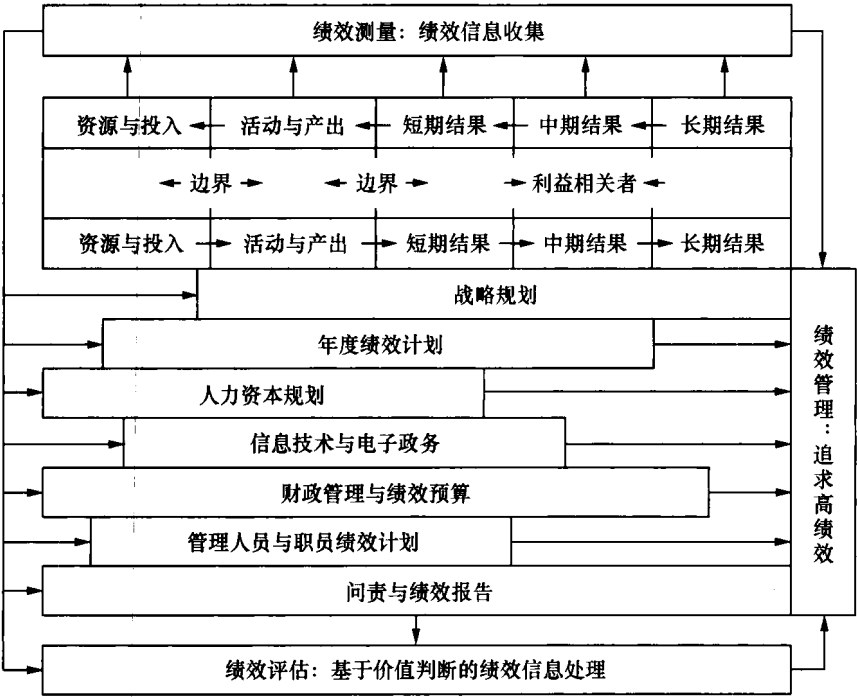


图 1 绩效“测量”、“评估”与“管理”的关系

2 公共部门绩效评估的理论框架

绩效评估是一个复杂的社会过程,涉及众多利益相关者。一定程度上,绩效评估是一种“平衡术”,其艺术性大于科学性,政治性多于技术性。从系统论的角度看,公共部门绩效评估的构成要素包括评估对象、评估定位、评估指标、评估组织者、评估主

体、绩效信息收集、评估结果使用等。上述要素之间并非相互独立,而是相互依存、紧密联系,共同构成公共部门绩效评估体系(如图 2 所示)。

绩效评估首先需要明确“评估什么”,也即明晰评估对象及其职能构成与特征、规模与数量等。评估对象的数量多少、规模大小、地域分布、职能与结构的复杂程度、可比性等特征都会对评估的复杂性、

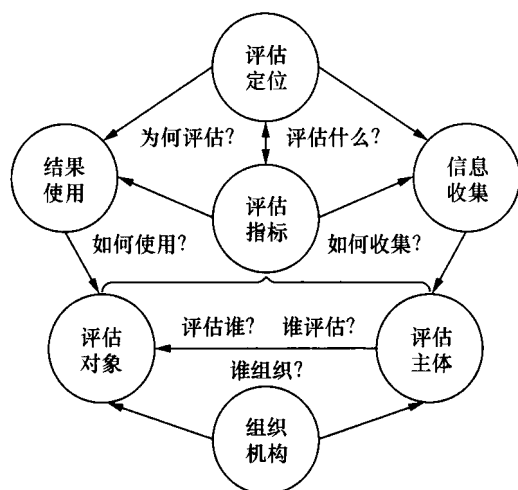


图2 公共部门绩效评估的构成要素及其相互关系

难度、成本和效果等产生深刻影响。公共部门绩效评估的对象有别于企业，缺乏可资评鉴的利润“底线”，因而难以进行衡量和比较，需要特别注意<sup>[2]</sup>。

评估定位或目的是绩效评估的意图和理由，即“为什么评估”，亦即评估活动的预计收益或结果。“思路决定出路，定位决定地位”。不同的评估定位在很大程度上决定了评估的效果和影响。从管理上看，评估定位包括改进、预算、庆祝、控制、问责、激励、学习<sup>[6]</sup>以及降低成本、创新等。其他评估定位还有：显在定位与潜在定位、短期定位与长期定位。现实的评估定位往往是复杂和微妙的，需要根据评估的动力和需求，综合多种定位而形成，很多时候“只可意会不可言传”。评估定位不同，所采用的评估指标等不同，给评估对象带来影响如压力等也不同，没有压力不行，压力过大也不行。因此，评估定位的选择应十分审慎。

评估指标反映的是特定评估对象的评估范围和内容，往往针对评估对象工作职能的某些方面或全部内容进行设计，并反映评估对象的绩效产生和形成机理<sup>[4]</sup>。评估指标的来源、设计和筛选是多渠道、多方面的，现实中需遵循特定原则，如 SMART（具体明确、可测量、可实现、相关性及时间性等<sup>[5]</sup>）。评估指标的权重具有调整 and 平衡的功能，使作为“指挥棒”的评估指标的导向作用得以实现，也是绩效评估中较难把握和处理的内容。权重设置的方法大致包括两大类：主观赋权法，如德尔菲法、层次分析法等，一般由专家根据经验对指标的重要性打分，所有分值汇总后设置权重；客观赋权法，如主成份法、神经网络法等，通常需要历史数据支撑，通过公式推导和统计分析得出。两种方法各有利弊，实际中往往结合使用。

在绩效评估的诸多利益相关者中，评估组织者的作用在于有效组织、管理、协调评估活动的开展，其基本职能包括评估方案的制定、评估主体的选取、评估过程的监督、评估的实施和反馈，以及整体工作的统筹协调。组织者是绩效评估的指挥所，也往往是绩效评估的信息汇集中心，其选择应充分考虑权力、技能、资源、成本、利益相关性等因素。根据委托-代理理论，评估委托方与代理方（组织者）之间存有多种委托代理关系，容易导致博弈、信息失真和误差放大等。评估组织者应区分决策与执行职能，兼顾决策权威性与执行灵活性。

绩效评估本质上是人的价值判断行为，评估指标的理解、评估程序的服从，往往因评估主体的不同而有差异，因此评估主体在绩效评估中占据重要地位<sup>[7]</sup>。绩效评估的客观性和准确性受到评估主体的信息掌握和知识拥有程度、利益相关程度等的影响<sup>[8,9]</sup>。评估主体的选择首先应考虑拥有相关信息和知识的主体，其次可以从评估对象涉及的利益相关者中选取（上级、同级、服务对象等），或者从与评估对象不存在利益关系的群体中选取（第三方独立），以及组成多元化的评估主体组合，如人力资源管理广泛采用的 360 度绩效评估。

绩效评估是一个绩效信息收集、处理、沟通、使用和反馈的过程。绩效信息是评估主体进行价值判断的依据，绩效信息是否充分、准确、客观，是决定评估有效性的关键因素。绩效信息来源包括自我报告、评估主体收集、第三方提供等。绩效信息的收集途径包括查阅文件、资料和报表、工作记录、现场考察、走访、座谈和重点访谈、问卷调查等。评估中需要综合运用，并考虑收集成本、信息的可靠性、收集效率等问题。

评估结果使用是绩效评估的最后环节，它是保证评估活动不流于形式，促进绩效不断改进的重要手段；它能形成适当的压力，产生一定的动力，促进绩效的不断改进。评估结果使用是激励和问责的主要形式，其方式包括公开公示、奖惩、配置资源、改进绩效、调整评估定位等。评估结果使用力度可强可弱，一般取决于评估定位和评估结果的信度与效度。

绩效评估是普遍行为，公共部门绩效评估的要素构成及其相互关系错综复杂，使评估体系呈现出复杂多样的特征，是否存在一种普适的理论框架可以对各种公共部门绩效评估活动进行分析和解释？本文认为，上述要素特征及相互关系，总括了公共部门绩效评估的体系框架，对于揭示绩效评估的规律

具有促进作用,可称为公共部门绩效评估的“元理论”。

### 3 公共部门绩效评估的“绩效”与创新

公共部门绩效评估本身也存在“绩效”高低问题,它既可以带来正面效果,也可能导致负面影响,是“双刃剑”。其正面效果如以评促建、强化监督与控制、增强利益相关者的信任和满意,实现组织良性发展等;负面效果则源自评估误差、实施走样及其他难以预料的影响,导致诸如利益相关者对评估冷漠、抱怨乃至消极抵触,组织功能紊乱和恶性循环等。

作为社会活动,公共部门绩效评估通常是利益调和与折衷的复杂过程,常常要反反复复、上上下下数次,有可能导致评估走样,但却是集体学习的理性选择结果<sup>[1]</sup>。因此,需要剖析绩效评估的前因后果,考察诸多要素对其绩效的影响。

公共部门绩效评估会随着社会和组织的发展而发生变化,并通过自身的创新来适应环境变化。而完善公共部门绩效评估的途径在于调整各个评估要素及相互关系,并对评估系统进行整体完善。评估要素及其相互关系是创新的素材和源泉,每个要素都有其变化的范围和方式,并可以通过调整和变动来实现其功能调节和创新。

评估对象创新。如将评估对象重新进行排列组合,按照不同类别进行评估;或者将适合评估的对象纳入评估框架,扩大评估范围等。

评估定位创新。如评估定位的更新、丰富和完善,如地方政府绩效评估从应付上级、GDP论英雄等转为科学发展观,从“为评估而评估”走向绩效改进。

评估指标创新。如引入以往关注不够的指标,如社会发展、环境保护、创新等;或者采用一套新的设计框架,如平衡计分卡、顾客满意度模型等。

评估组织者创新。如改变绩效评估的组织结构或变革评估组织者,如采用多级评估委员会制,或者为了体现公正性和无涉性,邀请第三方机构独立组织。

评估主体创新。如引入新的评估主体,如打破原来“唯上不唯下”、“对内不对外”的封闭评估系统,引入公众、第三方机构、专家学者等参与评估。

绩效信息收集机制创新。如采用新的信息收集方法,如网络评估、电话调查、绩效管理软件等,加快信息收集和处理速度、降低信息成本并扩大信息源。

评估结果使用方式创新。如扩大评估结果的使

用范围或强化使用力度,如过去多是评与不评一个样,评好评差一个样,现在则加大评估结果使用力度;而且使用方式也逐渐丰富,包括社会公开、晋升激励、问责、绩效预算等。

创新可以是单个要素的调整,有时则是多个要素的协同变动,甚至可能是全部要素的结构重构和革命。按照熊彼特的观点,创新本质上是一种“新的组合”,而上述绩效评估创新之间可以形成某种组合关系,进而产生不同的创新。

创新常常是组织内外两方面因素共同作用的结果,创新的动因既存在于客观的制度环境之内,也存在于主观的内在需求之中<sup>[10]</sup>。因此,创新型国家的建设需求和决策者的推动等都是公共部门绩效评估体系创新的动因。但公共部门绩效评估体系创新也是一把“双刃剑”,操作不当也会产生负面影响。尤须注意的是,它可能导致利益格局调整,并面临巨大的阻力,需要调和与满足关键利益相关者的需求。

### 4 公共部门绩效评估实践分析——以科技计划绩效评估为例

下文以科技计划绩效评估为例,对具有典型性的公共部门绩效评估实践进行简要分析。分析表明,尽管公共部门绩效评估理论框架可能无法完全回答绩效评估的所有问题,但它是目前较有系统性和包容性的理论框架,能够解释公共部门绩效评估的主要现象和关键问题,对于指导公共部门绩效评估的设计与实施具有一定价值。

建设创新型国家要求国家加大科技投入,而随着科技领域财政投入的加大,如何管好用好科技经费并发挥其积极效果,成为科技计划绩效评估的主要目标。2007年,财政部出台《中央级民口科技计划(基金)经费绩效考评管理暂行办法》,将科技计划/项目作为财政支出绩效评估的重要领域,完善绩效管理机制,提高经费使用效益。与此同时,地方政府和其他科研资助机构也纷纷出台管理措施,加强科技计划和项目的绩效评估。如国家自然科学基金委员会管理科学部从1998年开始就组织了对资助项目的后评估,取得了较好的效果。

科技计划绩效评估的对象是由政府资助的科技计划、科学基金、科技项目和专项课题,一般分为基础研究、应用研究、技术与开发等类别进行分类评估。科技计划绩效评估的定位是进一步强化管理控制,形成财政问责机制,提高资金使用效益。通过对科技计划的规划、实施和结果进行监测和评估,强

化财政投入部门和计划承担机构的责任意识,促进财政资金高效利用,并为未来财政支出乃至绩效预算提供参考依据。评估组织者多为财政管理和科技管理部门,一般通过招投标方式委托专业评估机构具体实施。评估主体主要为相关领域的专家,以保证其科学性和专业性。

科技计划绩效评估的重点是科技计划绩效目标的实现程度及其经费执行结果,从投入、活动与产出、结果与影响等方面综合衡量,评估科技计划的实际履行情况和综合效果。为了体现不同类型计划的差异,分别设置共性评估指标和根据计划特点确定的个性评估指标。评估通常由科技计划承担者提供数据,或者采用独立数据来源,分别指自报指标和源生指标<sup>[11]</sup>。一般在计划执行中期和终期进行评估,采取同行匿名评审、召开评审会或实地考察等形式。评估结果会直接影响计划承担者今后能否获得资助,并作为奖优罚劣的依据,发挥竞争性激励的正面效果。

科技计划绩效评估打破了重投入不重产出、重支出不重结果的传统科技财政管理模式,引入了问责和竞争机制,保障了科技计划财政支出结构和效益的优化。但由于尚处于起步阶段,科技计划绩效评估还存在一些亟待完善之处。评估对象是数量庞大、构成复杂的科技计划或项目,可比性低,评估难度大;科技含量高,往往“外行评内行”,评估结果可能失准。评估的强制性和约束力不够,科技计划主管部门参与评估的积极性不足,存在应付、走过场、形式化的问题。评估指标主要是量化数据,对科技计划的“潜绩”考量不足,通常难以反映真实效果和影响。此外,评估活动牵扯许多利益相关者,如财政部门、科技计划主管部门、评估机构、科技计划承担者等,存在多级委托-代理(principle-agent)关系,并可能干扰相关各方的日常工作。

完善科技计划绩效评估的可从以下几个方面入手:首先,建设动态的科技成果数据库,为科技计划绩效评估提供客观准确的数据来源;扶持和促进专业评估机构的发展,并形成合理有序的竞争态势,为科技计划绩效评估提供更多的选择。其次,在评估结果使用上,通过正面激励增强主管部门的评估动力,分步骤强化评估力度,实现评估的常态化、透明化以及制度化。最后,评估组织者可广泛开展试点与示范,通过绩效评估与标杆管理等促进优秀经验和最佳实践的推广,为科技计划绩效评估积累实践经验。

## 5 结束语

结合我国公共部门绩效评估的理论进展与实践状况,本文在综述绩效、绩效测量、绩效评估与绩效管理的概念异同与相互关系的基础上,提出了分析公共部门绩效评估的初步理论框架,包括公共部门绩效评估的七大要素,分别是评估对象、评估定位、评估指标、评估组织者、评估主体、绩效信息收集、评估结果使用等。上述要素相互联系,共同构成了公共部门绩效评估体系,形成设计、实施和完善公共部门绩效评估方案的基本出发点。论文进一步以科技计划绩效评估为例,说明了该理论框架的适用性,并指出其未来完善思路。

在全球化与信息化的背景下,各国政府都在追求高绩效管理,绩效评估是其手段之一。然而绩效评估是一个复杂、敏感的问题,往往“牵一发而动全身”,每每令研究者如履薄冰,需要与管理者齐心协力,循序渐进,共同推动公共部门绩效评估事业发展。对于管理者来说,尤其需要认识到公共部门绩效评估的“两面性”,不能一味推崇和盲目追求绩效评估,而应采取系统观的视角,审慎试点、稳步推进,真正发挥绩效评估的正面效果。

事实上,绩效评估适应了公共管理体系中组织间合作和系统治理的需求,使以往“单兵作战”的负责财政预算、人事任免和效能监察等各公共组织逐步加强合作,使下一盘“大棋”的可能性日益增加;作为人类的一种基本行为,绩效评估针对基本治理问题的反复讨论和碰撞,可能会成为各个领域思想与实践交流和融合的交叉点。

## 参 考 文 献

- [1] 吴建南, 阎波. 地方政府绩效评估体系的路径选择——福建的分析. 中国行政管理, 2008, (2).
- [2] 吴建南, 阎波. 政府绩效: 理论诠释、实践分析与行动策略. 西安交通大学学报(社会科学版), 2004, (3).
- [3] Wu J, Hou Y. The Rise of Public Administration in China: Reflections on and Inspirations from Globalization. Minnowbrook III: The Future of Public Administration, Public Management and Public Service around the World: Lake Placid, New York, USA, 2008.
- [4] 吴建南, 杨宇谦, 阎波. 政府绩效评价: 指标设计与模式构建. 西安交通大学学报(社会科学版), 2007, (5).
- [5] 西奥多·H. 波伊斯特. 公共与非营利组织绩效考评: 方法与运用. 北京: 中国人民大学出版社, 2005.
- [6] Behn RD. Why measure performance? Different purposes require different measures. Public Administration Review, 2003, 63 (5): 586—606.
- [7] 吴建南, 阎波. 谁是“最佳”的价值判断者: 区县绩效评估机制的利益相关主体分析. 管理评论, 2006, (4).

- [8] 吴建南, 岳妮. 利益相关性是否影响评价结果客观性: 基于模拟实验的绩效评价主体选择研究. 管理评论, 2007, (3).
- [9] 吴建南, 李玲. 知识拥有状况如何影响评价结果准确性? ——面向地方政府绩效评价的模拟实验研究. 中国管理科学, 2005, S1.
- [10] 吴建南, 马亮, 杨宇谦. 中国地方政府创新的动因、特征与绩效——基于“中国地方政府创新奖”的多案例文本分析. 管理世界, 2007, (8): 43—51.
- [11] 梁立明, 梁立华, 孔繁士. 高等院校科研绩效评估的自报指标与源生指标. 科研管理, 1998, (6).

## PERFORMANCE EVALUATION IN PUBLIC SECTORS: THEORETICAL FRAMEWORK AND PRACTICE ANALYSIS

Wu Jiannan

(School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an 710049)

**Abstract** Based on the definition of performance evaluation and its comparison with performance measurement and performance management, this paper proposes a preliminary theoretical framework composed of key factors and their interplay to analyze the design and implementation of performance evaluation in the public sectors. The performance of performance evaluation as well as its innovation are also discussed. Antecedents and consequences of some representative public performance evaluation activities in China, S&T plan performance evaluation, are briefly discussed based on the application of the theoretical framework.

**Key words** performance evaluation, theoretical framework, performance measurement, performance management, public sectors

• 资料 • 信息 •

## 吴飙等揭示非线性周期系统中能隙孤子和布洛赫波的对应关系

周期系统是物理学研究的焦点问题之一,例如晶格周期势场中电子的运动是固体物理学的基础。伴随实验技术的发展,最近涌现了一批新型的人工周期系统,比如光晶格中的超冷原子气,波导阵列和由光诱导的光子晶格,等等。这些系统除了都具有周期性外还具有非线性。这样,这些周期系统既可以拥有熟知的延展的布洛赫态又可以拥有局域的孤子解。这种在周期系统中的孤子解通常叫能隙孤子(gap soliton)。它分为基本(fundamental)能隙孤子和高阶能隙孤子。

最近,中国科学院物理研究所/北京凝聚态物理国家实验室凝聚态理论与材料计算室的研究员吴飙和博士生张永平揭示了一维非线性周期系统中这两种性质完全不同的解即布洛赫波和基本能隙孤子间存在着一种对应关系:布洛赫波可被看作为无数基本能隙孤子组成的孤子串。以此对应关系为基础,从布洛赫能谱出发,他们发现在不做任何计算的情

况下可以得到一系列基本孤子的一般性质。比如,对于排斥非线性,第 $n$ 个线性布洛赫能隙中存在 $n$ 族不同类型的基本孤子。以前人们仅知道前两族的存在,并错误地认为第二族和第一族不具有同等地位,把它叫做亚基本(sub-fundamental)孤子。吴飙和张永平进一步发现已知的许多这种非线性周期系统中的定态解包括高阶能隙孤子,能隙波,布洛赫波,多倍周期延展态都可以被看作由基本能隙孤子构成。这些理论结果对进一步系统而深入理解这些新的非线性周期系统具有重要意义。

上述工作结果发表在美国《物理评论快报》(Physical Review Letters (2009, 102: 093905))。相关研究得到了国家自然科学基金和“973”项目的支持。

(中国科学院 物理研究所 供稿)