



综述

可计算建模

汲培文^①, 江松^②, 张平文^{③*}

① 国家自然科学基金委员会数理学部, 北京 100085;

② 北京应用物理与计算数学研究所, 北京 100094;

③ 北京大学数学科学学院, 北京 100871

E-mail: jipw@nsfc.gov.cn, jiang@iapcm.ac.cn, pzhang@pku.edu.cn

收稿日期: 2012-02-24; 接受日期: 2012-04-13; * 通信作者

摘要 可计算建模 (computable modeling) 指根据所研究问题对计算精度的要求, 综合运用相关领域知识建立或简化模型, 减少计算量, 提高计算效率, 使得模型在现有计算机条件下可计算. 可计算建模是科学与工程计算研究的一个重要方面. 本文主要通过若干例子介绍可计算建模研究的内涵.

关键词 科学计算 可计算建模 共性算法

MSC (2010) 主题分类 65Z05, 65Z99

1 引言

科学计算是指利用应用数学和计算机科学所提供的计算能力来理解和解决科学与工程领域中的问题, 涉及建模、算法、程序、模拟和分析等过程. 科学计算是 20 世纪重要科学技术进步之一, 伴随着电子计算机的出现迅速发展并得到广泛应用. 科学计算已与理论研究和实验研究相并列成为科学研究的第三种方法, 成为促进重大科学发现和科技进步的重要手段. 国家重大战略需求中许多科学问题的解决高度依赖于科学计算中共性算法与可计算建模的发展水平.

共性算法是指通过计算模型构造出来的、经过理论分析和应用验证的普适性算法. 可计算建模指根据所研究问题对计算精度的要求, 综合运用相关领域知识建立或简化模型, 减少计算量, 提高计算效率, 使得模型在现有计算机条件下可计算.

模型与算法是计算数学研究的核心, 共性算法与可计算建模方法是每一个计算数学工作者追求的目标. 对共性算法的内涵大家认识比较一致, 如有限元方法、快速 Fourier 方法、快速多极子算法等就是典型的共性算法. 可计算建模是一个新的术语, 是在 2010~2011 年组织自然科学基金委员会重大研究计划“高性能科学计算的基础算法与可计算建模”过程中, 集大家智慧提出来的. 作为一个新的术语, 它的内涵是什么并不是很明确, 本文希望通过若干例子介绍可计算建模.

科学计算建模发展的一个趋势是由基于机理建模、数据建模到机理数据混合建模并存. 众多前沿科学领域中的重要实际问题常常涉及多个发生在不同时空尺度上相互耦合的物理过程, 具有高度的各向异性、奇异性、超高维、非均匀性以及不确定性等特征, 它们的建模十分复杂和困难. 如何根据这些复杂的实际问题对精度的要求和现有计算机的能力, 对其建立可计算的模型 (可计算建模) 以充分发挥计算机的能力和算法的特点是日前科学与工程计算面临的主要挑战和共性难题之一. 可计算建模研

究大致可以分为三个方面: 即模型约化, 多尺度多过程耦合建模与数学表达, 机理与数据混合建模. 下面将以若干例子介绍、分析此三方面的内涵.

2 模型约化

模型约化指根据所研究问题对计算精度的要求, 综合运用相关知识简化模型, 减少计算量, 提高计算效率.

2.1 密度泛函理论

波函数和 Schrödinger 方程在量子力学中起着十分重要的作用. 对一个给定的微观体系, 我们可能得到的所有信息都包含在系统的波函数中. 以一个外势场 $v(\mathbf{r})$ 中的 N 个电子体系为例, 量子力学的波动力学范式可以表示成

$$v(\mathbf{r}) \implies \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \implies \text{可观测量}.$$

上述过程即为: 对给定的外势, 求解 Schrödinger 方程可以得到电子波函数, 通过波函数计算力学量算符的期望值可以进一步得到所有可观测量物理量的值.

N 个电子体系的 Schrödinger 方程是 $3N$ 维的线性特征值问题, 当 $N > 2$ 时直接求解基本上不可能. 然而, 在大多数情况下, 我们只关心一些特定的物理量如能量、电子密度等. 所以人们希望使用一些比波函数更简单的物理量来求解, 如单粒子 Green 函数、密度矩阵、电子密度等. 从波函数到单粒子 Green 函数, 从单粒子 Green 函数到密度矩阵, 从密度矩阵到电子密度都丢失了大量信息, 所以按道理我们将不能从电子密度中“恢复”波函数. 但令人吃惊的是, 密度泛函理论给出了完全相反的结论, 至少对于只考虑基态的情形, 从波函数、Green 函数、密度矩阵到密度, 没有丢失任何信息. 电子密度不仅仅是众多可观测量中的一个, 而且是可以用来计算其他所有可观测量的体系基本物理量. 电子密度只是空间坐标的函数, 从而密度泛函理论将求解 $3N$ 维波函数的问题转化为求解 3 维电子密度的问题.

现代密度泛函理论的基础建立在 1964 年 Hohenberg 和 Kohn^[1] 提出的两个著名定理之上, 其中定理 1 指出, 在相差任意常数意义下, 不计自旋的全同 Fermi 子系统非简并基态的电子密度函数 n 唯一地决定外势 v ; 定理 2 给出了密度泛函理论的变分法. 为了将密度泛函理论实用化, 需要寻找能量泛函的具体表达形式, Kohn 和 Sham^[2] 在 1965 年提出了一个可能的方案, 他们引进了一个与相互作用多电子体系有相同电子密度的假想的非相互作用多电子体系, 从而得到著名的 Kohn-Sham 方程,

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) \right) \phi_i = \epsilon_i \phi_i,$$

其中 $v_{\text{ext}}(\mathbf{r})$ 、 $v_H(\mathbf{r})$ 、 $v_{xc}(\mathbf{r})$ 分别为外势、Hartree 势和交换关联势. 在 Kohn-Sham 方程中, 有效势 $v_{\text{eff}} = v_{\text{ext}} + v_H + v_{xc}$ 由电子密度决定, 而电子密度又由方程的特征函数求得, 所以需要自洽求解 Kohn-Sham 方程. 这种自洽求解过程通常称为自洽场方法.

在 Kohn-Sham 方程中, 起关键作用的是交换关联势 v_{xc} . 对于给定的系统, 交换关联势在原则上包含了系统基态的全部信息. 然而交换关联势是电子密度的复杂泛函, 其严格的具体形式目前并不知道. Hohenberg 和 Kohn 的理论也仅陈述了交换关联势的存在性, 而未涉及其构造的任何细节. 为了在实际中求解 Kohn-Sham 方程, 交换关联势的形式必须被简化. 目前广泛应用的交换关联势是基于

均匀电子气的交换关联势的逼近, 包括局域密度近似 (LDA) [3,4], 广义梯度近似 (GGA) [5-7], 以及杂泛函 [8] 等. 基于这些实际中可操作的交换关联势模型, 密度泛函理论成为一种完全基于量子力学的第一原理 (ab-initio) 理论, 将 $3N$ 维线性偏微分方程特征值问题简化为 3 维非线性偏微分方程特征值问题, 从而大大地减少了计算量, 使得许多量子化学问题的计算得以实现. 如今密度泛函理论已经成为量子化学和材料科学中应用广泛的理论, Kohn 也因此项工作获得了 1998 年度诺贝尔化学奖. 可以说密度泛函理论是可计算建模方法研究中最成功的范例.

密度泛函理论的广泛应用还和赝势理论的发展有着密切的联系. 虽然 Kohn-Sham 方程是一个 3 维的特征值问题, 原子核产生的库仑外势导致 Kohn-Sham 方程的特征函数在原子核附近光滑性不好. 为了描述发散的库仑外势和不光滑的特征函数, 计算网格必须充分稠密. 这使得求解 3 维 Kohn-Sham 方程的计算量依然很大, 从而限制了密度泛函理论在实际中的应用. 然而在一个化学系统中, 所有的电子起的作用并不都是等价的. 电子分价电子和核电子两种, 其中核电子不参与化学反应, 并总是和原子核绑定在一起, 只有价电子参与化学成键和化学反应. 赝势理论基于这个观察对于全电子系统的计算进行了简化, 将原子核和核电子作为一个整体看待, 而用原子核与核电子共同产生的有效外势求解只含有有效价电子的 Kohn-Sham 方程. 在赝势理论中, 有效外势在整个空间中是有界的, 而对应的有效价电子的特征函数是光滑的, 从而使用赝势的 Kohn-Sham 方程可以以比全电子的 Kohn-Sham 方程低得多的代价求解. 目前广泛适用的赝势有模守恒赝势 [9,10]、超软赝势 [11] 等. 赝势理论综合了对化学系统已有的了解, 简化了计算量, 使得密度泛函理论在 80 年代已经具备处理实际体系的能力, 是可计算建模的典型例子.

密度泛函理论的传统应用是计算有序晶体和小分子的能量和结构. 近二十年来, 随着密度泛函理论不断发展, 密度泛函理论被越来越多地应用在大规模、非周期的无序系统上, 如材料缺陷、复杂的化学反应、生物大分子的模拟等. Kohn-Sham 方程特征值求解的计算量与系统电子数目的三次方成正比, 这使得密度泛函理论的一般应用范围只能局限在 100 至 1000 个原子的尺度. 为了实现 1000 至 10000 个原子以上尺度的第一原理计算, Kohn-Sham 方程的特征值求解过程需要进一步地简化. 线性标度方法 [12] 是一类对绝缘体和半导体适用的有效方法. 线性标度法的基础是电子的局域性原理, 其含义是在绝缘体和半导体系统中, 系统外势的局域扰动对远处电子的影响是指数衰减的 [13]. 局域性原理使得大系统的求解可以被分解为许多小问题, 每个小问题可以分别求解, 从而整体计算量是系统原子数目的线性函数. 但是, 局域性原理并不是对所有系统普适的原理. 在金属系统中, 系统外势的局域扰动对远处电子的影响是多项式衰减的, 因而线性标度法也不能被有效地应用在金属系统的计算中. 如何设计适用于包括金属、半导体和绝缘体在内的各种体系的有效共性算法, 实现大规模系统的可计算建模, 是计算数学、化学和材料科学共同面临的难题.

2.2 自洽平均场理论

自洽平均场理论是一个在不同时期、地点, 以不同的形式, 被不同的科学家重复发现的理论. 其最早的起源可以追溯到 1873 年 Van der Waals [14] 提出的气液状态方程, 他在研究气液相变时, 引入了“内压力”来描述气体分子间在较远距离上的相互作用, 从而修正了理想气体的状态方程. 平均场理论的核心思想是引入一个等效场 w , 将一个多体作用体系变成在等效场 w 作用下的单体体系, 将基于粒子的体系变成基于场的体系. 此等效场 w 是由其他粒子与某个指定粒子下相互作用的平均. 假设系统可以用密度 ρ 这个序参量来描述, 所谓“自洽”, 就是密度 ρ 可以决定等效场 w , 而反过来 w 又能决定 ρ , 这样相互制约, 相互作用, 从而达到平衡, 这就是自洽. 一般而言, 自洽平均场方程可以表示成以下

形式:

$$w = w[\rho], \quad (2.1)$$

$$\rho = \rho[w]. \quad (2.2)$$

方程 (2.1) 一般来自对多体相互作用能量函数的分析, 方程 (2.2) 则来自对单体问题本身的分析. 从数学上, 这种建模的方式将构造出一个非线性变分问题. 目前的平均场理论仍然是在平衡态统计力学的框架下, 对于平衡态静力学能比较好地处理. 也就是说引入的等效场 w 是在某种平均意义下的场形式 \bar{w} , 而密度算子也是某种平均 $\bar{\rho}$. 这样系统中的每个单体都是等价的.

下面通过一个例子来说明目前平均场理论比较通用的建模流程. 考虑一个具有 n 个相同粒子的多体作用体系, 其每个粒子的空间位置为 $\hat{R}_i, i = 1, 2, \dots, n$, 其中 $\hat{\cdot}$ 表示微型的物理量. 我们只考虑相互作用的能量, 且只与粒子的空间位置有关, 在微观状态 $\hat{\mathbf{R}} = (\hat{R}_1, \hat{R}_2, \dots, \hat{R}_n)$ 下, 系统的能量表示为

$$U(\hat{\mathbf{R}}) = U(\{\hat{R}_i\}_{i=0}^n) = \sum_{i < j} u(\hat{R}_i, \hat{R}_j). \quad (2.3)$$

体系在空间 \mathbf{r} 处的数密度为

$$\hat{\rho}(\mathbf{r}) = \sum_{i=1}^n \delta(\mathbf{r} - \hat{R}_i). \quad (2.4)$$

将数密度表达式代入到系统能量方程得

$$U(\hat{\mathbf{R}}) = \int d\mathbf{r} \int d\mathbf{r}' \hat{\rho}(\mathbf{r}) u(\mathbf{r}, \mathbf{r}') \hat{\rho}(\mathbf{r}') \triangleq U[\hat{\rho}]. \quad (2.5)$$

基于正则系综^[15], 系统的配分函数可表述为

$$Z = \prod_{i=1}^n \int d\hat{r}_i e^{-\beta U[\hat{\rho}]}, \quad (2.6)$$

其中 $\beta = 1/(k_B T)$, k_B 是 Boltzmann 常数, T 为温度. 引入 δ 泛函

$$\int D\rho \delta[\rho - \hat{\rho}] = 1, \quad (2.7)$$

将其代入到配分函数中, 则有

$$Z = \prod_{i=1}^n \int d\hat{R}_i \int D\rho e^{-\beta U[\rho]} \delta[\rho - \hat{\rho}]. \quad (2.8)$$

将 δ 泛函表示成 Fourier 形式,

$$\delta[\rho - \hat{\rho}] = \int_i Dw \exp \left\{ \int d\mathbf{r} w(\mathbf{r}) [\rho(\mathbf{r}) - \hat{\rho}(\mathbf{r})] \right\}, \quad (2.9)$$

其中 \int_i 表示沿虚轴积分, 从而形式上将基于粒子的配分函数变成了基于场的函数

$$Z = \int D\rho \int_i Dw \exp \left(-\beta U[\rho] + \int d\mathbf{r} w \rho \right) \prod_{i=1}^n \int d\hat{R}_i \exp \left(- \int d\mathbf{r} w \hat{\rho} \right)$$

$$= \int D\rho \int_i Dw \exp\left(-\beta U[\rho] + \int d\mathbf{r} w\rho\right) (Q[w])^n, \quad (2.10)$$

其中

$$Q[w] = \int d\hat{R}_i \exp\left(-\int d\mathbf{r} w\hat{\rho}\right) \quad (2.11)$$

是一个粒子在场 w 作用下的单粒子配分函数. 因为在平均场理论中, 同一种性质粒子的地位一样, 所以每个粒子的状态都是独立同分布. 根据系综理论,

$$\bar{\rho}_i = -\frac{\delta \log Q[w]}{\delta w}. \quad (2.12)$$

将单粒子配分函数代入到配分函数 (2.10), 得到

$$Z = \int D\rho \int Dw \exp(-H[\rho, w]), \quad (2.13)$$

其中 Hamilton 量为

$$H[\rho, w] = \beta U[\rho] - \int d\mathbf{r} w\rho - n \log Q[w]. \quad (2.14)$$

以上的推导成功地将多体相互作用体系转化为非线性变分问题, 并且没有做任何近似. 但注意到配分函数 (2.13) 是一个泛函积分, 在绝大多数情况下是难于处理的^[16], 因此需要近似. 鞍点近似 (也称为平均场近似) 就是其中的一种. 鞍点近似是只考虑最大几率的平衡态, 同时忽略平衡态局部热涨落的一种近似^[17]. 利用鞍点近似, 配分函数 (2.13) 可近似为

$$Z \sim H[\bar{\rho}, \bar{w}]. \quad (2.15)$$

因此 $H[\rho, w]$ 在平衡态 $(\bar{\rho}, \bar{w})$ 处的一阶变分为零, 得到平均场方程

$$\beta \frac{\delta U[\bar{\rho}]}{\delta \bar{\rho}} - \bar{w} = 0, \quad (2.16)$$

$$n \left(\frac{\delta \log Q[\bar{w}]}{\delta \bar{w}} \right) + \bar{\rho} = 0. \quad (2.17)$$

这个例子针对的是只含有同一种性质粒子的体系, 当体系中还有 m 种不同性质的粒子时, 就会出现 m 个不同的单粒子配分函数. 这时候得到的平均场方程可以处理更为复杂的情况, 例如体系中含有结晶的和非结晶的两类高分子. 自洽平均场理论将多体相互作用体系转化为非线性变分问题, 从而可以减少计算量, 大幅提高计算效率.

平均场理论是伴随着连续相变的研究而发展起来的, 在早期的气液固三态的相变, 以及铁磁相变、超导相变、超流相变、合金的有序-无序相变等领域已有非常广泛的应用. 近年来, 平均场理论也被广泛地运用到研究软物质凝聚态体系的相行为, 包括寻找自组装所形成的有序结构, 构造相图, 研究不同结构之间的相变过程 (包括无序到有序, 有序到有序), 以及不同结构之间的界面形成等. 但平均场模型通常是一类非线性的、具有多解、多参数的变分问题, 目前在求解上还存在不少需要克服的问题. 通常而言, 数值求解面临的难题主要体现在三个部分: 离散格式、迭代格式和初值的筛选. 其中迭代格式严重地依赖于模型所提供的信息, 因此在构造迭代格式的时候需要对模型的数学性质进行很好的分析, 必要时需要在相应的理论框架下重新构造具有良好数学性质的物理模型. 初值的选取则依赖于对问题物理机制的理解, 可采用相应的数学理论或特殊的技术来获得. 例如在寻找嵌段共聚体系自组装形成的周期结构时, 在理解模型参数的物理意义下, 可采用空间群理论来筛选初值^[18].

2.3 矩封闭方法

矩封闭方法是一类常用的模型约化方法, 例如流体力学 Navier-Stokes 方程由流体的质量守恒、动量守恒和能量守恒导出, 但得到的方程不封闭. 流体力学方程封闭需要本构关系, 如假设应力是应变的线性函数 (Newton 流). 但对一些复杂的流体, 如生物流体, 本构方程不再是线性的, 人们把这类流体称为非 Newton 流^[19]. 下面我们以高分子稀溶液的小球弹簧模型为例, 来阐明矩封闭方法^[20,21].

高分子稀溶液的小球弹簧模型可以表述为宏观的不可压缩 Navier-Stokes 方程

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p &= \nabla \cdot \boldsymbol{\tau}_p + \nu \Delta \mathbf{u}, \\ \nabla \cdot \mathbf{u} &= 0,\end{aligned}$$

耦合微观的 Fokker-Planck 方程

$$\frac{\partial f}{\partial t} + \mathbf{u} \cdot \nabla_x f + \nabla_Q \cdot (\boldsymbol{\kappa} \cdot \mathbf{Q} f) = \frac{2}{\zeta} \nabla_Q \cdot (\mathbf{F}^c f) + \frac{2}{\zeta} \Delta_Q f,$$

其中 x 是空间点, \mathbf{Q} 是构形空间向量, $\mathbf{u} = \mathbf{u}(x, t)$ 表示流体的速度 $f = f(x, \mathbf{Q}, t)$ 表示构形分布函数, p 表示压力, ν 表示背景流体的粘性系数, $\boldsymbol{\kappa} = (\nabla_x \mathbf{u})^T$ 为速度梯度, $\boldsymbol{\tau}_p$ 表示高分子的应力, 可以通过构形分布函数表示:

$$\boldsymbol{\tau}_p = \langle \mathbf{Q} \nabla_Q \mu \rangle = -n \mathbf{I} + n \langle \mathbf{Q} \mathbf{F}^c \rangle,$$

这里 $\langle \cdot \rangle$ 表示构形分布函数的平均. $\mathbf{F}^c = \nabla_Q U$ 为弹簧力, U 为弹簧势能, 其中 Hookean 弹簧

$$U = \frac{1}{2} H |\mathbf{Q}|^2,$$

FENE 弹簧

$$U = -\frac{1}{2} H \log(1 - |\mathbf{Q}|^2/b).$$

Navier-Stokes 方程是非线性三维偏微分方程, 构形分布函数满足的 Fokker-Planck 方程是六维偏微分方程. 矩封闭方法就是在一定条件和精度要求下, 约化 Fokker-Planck 方程. 高分子的应力可以用构形分布函数的某种平均表示, 对 Hookean 弹簧, 高分子的应力为构形分布函数二阶矩 $\mathbf{M} = \langle \mathbf{Q} \mathbf{Q} \rangle$, 从 Fokker-Planck 方程可以导出二阶矩方程, 这就是著名的 Oldroyd-B 模型. 对 FENE 弹簧, 由于此时弹簧力不再是线性的, 则二阶矩方程不封闭, 需要作一些假设才能封闭二阶矩方程.

矩封闭有很多方法, 一个好的方法需要考虑如下几个方面: 矩封闭后还满足能量耗散, 能够重构非负分布函数, 具有好的精度, 计算量大幅减少. Bingham 近似是一类非常好的矩封闭方法, FENE 弹簧模型的 Bingham 近似就是把构形分布函数投影到没有流场的 Fokker-Planck 方程的平衡态解空间, 具体过程如下.

构形分布函数表达为下面特殊形式:

$$f(\mathbf{Q}) = \frac{1}{Z(\mathbf{R})} \left[1 - \frac{|\mathbf{Q}|^2}{b} \right]^{H/2} \exp \{ \mathbf{R} : \mathbf{Q} \mathbf{Q} \},$$

这里 $Z(\mathbf{R})$ 是为了满足 $\int f d\mathbf{Q} = 1$, 定义如下:

$$Z(\mathbf{R}) = \int \left[1 - \frac{|\mathbf{Q}|^2}{b} \right]^{H/2} \exp \{ \mathbf{R} : \mathbf{Q} \mathbf{Q} \} d\mathbf{Q},$$

\mathbf{R} 是矩阵形式的 Lagrange 乘子, 由下面方程确定:

$$\mathbf{M} = \frac{1}{Z(\mathbf{R})} \int \mathbf{Q} \mathbf{Q} \left[1 - \frac{|\mathbf{Q}|^2}{b} \right]^{H/2} \exp \{ \mathbf{R} : \mathbf{Q} \mathbf{Q} \} d\mathbf{Q},$$

从而可以得到代替 Fokker-Planck 方程的二阶矩方程

$$\frac{\partial \mathbf{M}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{M} - \boldsymbol{\kappa} \cdot \mathbf{M} - \mathbf{M} \cdot \boldsymbol{\kappa}^T = -\frac{4}{\zeta} (\mathbf{R} \cdot \mathbf{M} + \mathbf{M} \cdot \mathbf{R}).$$

高分子的应力张量可以表达为

$$\boldsymbol{\tau}_p = n (\mathbf{R} \cdot \mathbf{M} + \mathbf{M} \cdot \mathbf{R}).$$

约化后的二阶矩方程是三维的非线性偏微分方程. 可以证明 Navier-Stokes 方程与上述二阶矩方程的耦合体系仍然满足能量耗散, 由于从六维问题简化为三维问题, 计算量大幅减少.

3 多尺度多过程耦合建模与数学表达

材料损伤与破坏、流体湍流、核爆炸过程、生物大分子等现象呈现出巨大的尺度效应, 并伴随着不同尺度上的物理多样性和强耦合性, 其结果通常呈现出问题具有高度的各向异性、非均匀性以及不确定性等特征. 这些特征不仅使得建模过程和模型渐趋复杂, 而且一般单一尺度模型往往难以精确刻画, 因而目前更倾向于使用多物理、多尺度以及多场耦合模型^[22], 也即在不同区域或不同物理层次应用根据不同物理规律建立的数学模型, 并将这些模型进行耦合求解. 这些现象在不同尺度不同层次的原理机制已经比较清楚, 但核心问题是如何进行跨尺度关联与多过程耦合, 其中的根本问题是建立不同层次的物理模型、物理过程之间的关系, 以及这些关系的数学表示, 例如如何从底层的微观模型 (如分子动力学模型或量子力学模型) 出发获得介观或连续尺度上的粗粒化数学模型. 多尺度算法的挑战是如何将多时空、多过程的模拟进行无缝耦合, 特别是动力学耦合, 以及研究具有相同内涵不同变量的相互表达和转换关系、耦合区域或界面的确定原则, 满足物理守恒律且易于计算的连接条件的数学表达等.

3.1 拟连续体方法与数学理论

拟连续体方法是当前固体研究中影响较大的一种多尺度方法^[23], 它已被用于模拟一系列晶态固体的基本问题, 其中包括断裂、晶界结构与变形、纳米压痕以及三维位错分析. 拟连续体方法是一种区域分裂型的方法, 它把求解区域分为局部区域和非局部区域, 其中非局部区域往往覆盖晶体缺陷, 而局部区域往往变形梯度较小. 局部区域和非局部区域通过自适应准则自动确定. 局部区域使用 Cauchy-Born 法则计算应变能^[24], 而非局部区域往往利用原子相互作用势函数来计算系统能量. 具体而言, 如果用 x 和 y 分别表示原子在变形前和变形后的位置, 并且视 y 为 x 的函数, 则系统的总能量可以表示为:

$$E^{\text{tot}}[y](x) = E^{\text{loc}}[y](x) + E^{\text{nonloc}}[y](x).$$

给定边界条件和外力, 极小化系统总能量 $E^{\text{tot}}[y](x)$ 来得到原子的构型. 当然, 势函数并不一定是经验势, 系统能量也可以从量子力学模型出发, 通过从头计算的方法得到. 同时为了减少系统的总自由度, 在计算过程中只选取远远少于原子总数的少数代表性原子 (representative atoms) 来计算系统能量, 而其他原子的各项参数通过这些代表原子的参数插值得到.

微观和宏观通过 Cauchy-Born 法则以及交界面应力的连续性进行耦合. 如果这些连续性条件得不到满足, 就会产生非物理的力, 即所谓的“鬼力”(ghost force) [25]. 从物理上来说, 鬼力会导致系统产生非物理的塑性形变, 从而给出错误的物理图象; 从数学上来说, 鬼力会导致算法不收敛. 文 [26] 中构造了一个一维的例子, 通过研究显式解, 严格证明了鬼力会导致方法不收敛. 目前已有几种无鬼力的拟连续体方法 [27, 28], 但是对于非平面的微宏观交界面情形, 目前尚无成功范例报道. 值得注意的是, 还有一种所谓基于力的拟连续体方法, 该方法直接耦合宏微观平衡方程而不是耦合宏微观能量. 与以上基于能量的方法相比, 这种方法没有鬼力的问题, 但是由于没有系统总能量, 这种方法的数值实现会有一些困难, 它可能收敛很慢或收敛到错误的解, 甚至不收敛, 需要设计特别的迭代算法来求解相应的非线性问题 [29]. 该方法已成功地应用于固体材料中的静态和拟静态问题的模拟, 但也有不少对于有限温度情形和动力学问题的有益尝试.

对于动力学问题而言, 除了需要处理静力学问题的各种困难外, 还有一个最大的挑战就是要面对不同区域之间的界面可能引起的波反射. 以异质多尺度 (多物理) 方法为例 [30], 其仍然将求解区域分成连续性区域和原子区域.

在原子区域采用分子动力学模型, 研究对象为每一个原子的位置 \mathbf{x}_i 和速度 \mathbf{v}_i , 描述原子之间相互作用的势函数 V 可以是经验势, 也可以通过量子模型计算得到:

$$\begin{cases} \dot{\mathbf{x}}_i = \mathbf{v}_i, \\ m_i \dot{\mathbf{v}}_i = -\nabla_{\mathbf{x}_i} V. \end{cases}$$

在连续性区域采用非线性弹性动力学, 研究对象为位移场 \mathbf{u} 、应力场 $\boldsymbol{\sigma}$ 和能量 e , 为了保证和原子模型的相容性, 其本构关系 $\boldsymbol{\sigma}$ 与 \mathbf{J} 由原子模型通过 Cauchy-Born 法则给定:

$$\begin{cases} \rho_0 \frac{\partial^2 \mathbf{u}}{\partial t^2} = \nabla \cdot \boldsymbol{\sigma}, \\ \rho_0 \frac{\partial e}{\partial t} = -\nabla \cdot \mathbf{J}. \end{cases}$$

原子区域与连续性区域之间的界面条件是在力学平衡态下对两种模型进行线性化得到的相容性格式:

$$m_i \ddot{\mathbf{x}}_i = -\nabla_{\mathbf{x}_i} V(\mathbf{x}_i, B_{ij} \mathbf{x}_i) - \int_0^{t-t_n} \theta(\tau) \tilde{\mathbf{v}}(t-\tau) d\tau - \theta(t-t_n) \tilde{\mathbf{x}}_i(t_n),$$

其中, 右端项的第一部分处理的是波长较长的光滑部分, 允许这样的波在原子区域和连续性区域之间来回自由地传播, 而第二部分则将原子区域向连续型区域传播的波长较短且频率较高的那些波过滤掉. 这里的 θ 就是原子模型的离散吸收边条件对应的核, 在实际计算的时候需要做进一步的近似.

该格式应该是建立在分子动力学变分边界条件上的一种推广, 从单纯的处理声子在独立原子区域边界上的反射扩展到原子模型和连续模型耦合时处理声子从原子模型区域到连续模型区域的反射, 同时还需要保证弹性波在两个模型之间的自由传播.

这样的方法已被成功地应用于脆性裂纹扩展的研究, 能非常好地处理裂纹扩展, 即尖端产生的弹性波与声子波向边界的传播和边界条件产生的弹性波向裂纹尖端的传播, 以及这两者之间的相互作用 [30].

3.2 输运与扩散耦合建模与数学表达

微观粒子在介质中的迁移现象广泛地存在于自然界, 如大气中的光散射, 核反应堆中的中子输运, 宇宙星体大气中的光辐射, 等离子体中带电粒子的传输等许多物理现象中, 都会出现粒子传输过程. 研

究大量微观粒子在介质中迁移统计规律的数学理论称为输运理论. 微观粒子可以是中子、电子、离子、分子和光子等. 输运理论在科学和工程的众多领域有着非常广泛的应用. 描述粒子迁移过程的输运方程, 实质是微观粒子在介质中传输的守恒关系表达式.

输运方程也称为粒子数守恒方程, 或者 Boltzmann 方程. 根据输运理论可以很容易地建立起输运方程, 但输运方程的数值求解非常困难, 主要原因在于输运方程的求解计算量十分巨大 (如三维非定常输运方程的数值模拟需要求解七维问题), 当前的计算机能力还不能满足要求. 在一些条件下, 可以用能满足实际精度要求的简化模型来处理输运过程, 以大大减少计算量, 使得输运问题的数值求解在现有的计算机条件下得以实现. 下面以光子的输运过程即辐射输运为例, 简单介绍输运过程的输运与扩散耦合物理建模, 它既能满足某些实际问题对数值模拟精度的要求, 又能大大地减少计算量.

辐射输运理论的基本出发点是光的粒子性, 采用粒子观点, 建立光子分布函数的动理学方程. 辐射在介质中传播, 会发生吸收、发射以及散射等物理过程, 从而和物质之间交换动量和能量. 辐射输运是在介质中通过光子的发射和吸收而实现的能量重新分布过程, 物质的温度和密度决定辐射场的吸收和发射, 同时辐射场又影响物质的温度和密度及其运动状态.

辐射输运方程实质是光子数守恒的数学描述, 物质处在局部热动平衡下辐射输运方程的一般形式如下 [31, 32]:

$$\frac{1}{c} \frac{\partial I(\nu, \vec{\Omega})}{\partial t} + \vec{\Omega} \cdot \nabla I(\nu, \vec{\Omega}) = \sigma'_a(\nu) [B(\nu) - I(\nu, \vec{\Omega})] + \int_0^\infty d\nu' \int_{4\pi} d\vec{\Omega}' \frac{\nu}{\nu'} \sigma_s(\nu' \rightarrow \nu, \vec{\Omega}' \cdot \vec{\Omega}) I(\nu', \vec{\Omega}'), \quad (3.1)$$

其中 $I(\nu, \vec{\Omega})$ 是谱辐射强度, ν 是辐射频率, $\vec{\Omega}$ 是辐射传输方向, $\sigma'_a(\nu)$ 是有效吸收系数, $\sigma_s(\nu' \rightarrow \nu, \vec{\Omega}')$ 是散射系数, $B(\nu)$ 是 Planck 黑体辐射谱:

$$B(\nu) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/(kT)} - 1}, \quad (3.2)$$

h 为 Planck 常数, k 为 Boltzmann 常数.

谱辐射强度 $I(\nu, \vec{\Omega})$ 实质上依赖七个自变量 (频率 ν 、空间方向 $\vec{\Omega}$ 、空间位置 \vec{r} 和时间 t). 如前所述, 一般情况下, 输运方程的求解计算量巨大, 很难求解. 对于非平衡辐射场情形, 对辐射的频率依赖需要仔细考虑, 可以采用多群的方法描述. 当辐射场是弱各向异性时, 谱辐射强度对角度的依赖性很弱, 可以进行球谐函数展开, 考虑前两项近似 (称为 $P-1$ 近似):

$$I(\nu, \vec{\Omega}) = \frac{1}{4\pi} I_0(\nu) + \frac{3}{4\pi} \vec{\Omega} \cdot \mathbf{I}_1(\nu), \quad (3.3)$$

其中

$$I_0(\nu) = \int_{4\pi} d\vec{\Omega} I(\nu, \vec{\Omega}), \quad \mathbf{I}_1(\nu) = \int_{4\pi} d\vec{\Omega} \vec{\Omega} I(\nu, \vec{\Omega}). \quad (3.4)$$

为简单起见, 下面讨论忽略光的散射过程. 分别将方程 (3.1) 两边对立体角积分和乘以 $\vec{\Omega}$ 后对空间立体角积分, 得到

$$\begin{cases} \frac{1}{c} \frac{\partial I_0(\nu)}{\partial t} + \nabla \cdot \mathbf{I}_1(\nu) = \sigma'_a(\nu) [4\pi B(\nu) - I_0(\nu)], \\ \frac{1}{c} \frac{\partial \mathbf{I}_1(\nu)}{\partial t} + \frac{1}{3} \nabla I_0(\nu) = -\sigma'_a(\nu) \mathbf{I}_1(\nu). \end{cases} \quad (3.5)$$

进一步, 当 $\frac{1}{c} \frac{\partial \mathbf{I}_1(\nu)}{\partial t}$ 很小时, 即可得到扩散方程:

$$\frac{1}{c} \frac{\partial I_0(\nu)}{\partial t} + \nabla \cdot \mathbf{I}_1(\nu) = \sigma'_a(\nu) [4\pi B(\nu) - I_0(\nu)], \quad (3.6)$$

其中

$$I_1(\nu) = -\frac{1}{3\sigma'_a(\nu)} \nabla I_0(\nu).$$

有效吸收系数的倒数是辐射自由程, 它是辐射与物质相互作用强弱的度量. 当辐射的平均自由程远小于物质系统的尺度时, 物质是光性厚的, 在传输过程中, 辐射要经过多次的吸收和发射. 在这种情况下, 用扩散模型 (3.6) 描述输运过程原则上是精确的. 相对于输运模型 (3.1), 扩散模型 (3.6) 大大地减少了计算量 (对三维问题, 自变量为五个). 当辐射的平均自由程大于系统的尺度时, 系统是光性薄的, 在传输过程中只有一部分辐射要被吸收和发射, 需要用输运模型. 在一些实际输运问题中, 某些区域是光性厚的, 可以用扩散模型模拟, 余下的区域用输运模型模拟, 即输运扩散耦合模型. 与输运模型相比, 它的数值求解可大大地节省计算量. 例如在激光间接驱动惯性约束聚变中, 激光入射到高 Z 的圆柱腔 (黑腔), 黑腔内填充低 Z 气体, 中心是装有 DT 的内爆靶丸. 激光在黑腔壁被吸收转换成 X 射线辐射, 通过辐射输运过程, 实现为靶丸内爆创造一个合适的辐射场环境, 驱动靶丸形成内爆压缩达到 DT 点火, 实现热核放能.

在激光聚变数值模拟方面, 黑腔腔内是低密度的低 Z 等离子体, 辐射自由程长, 是光性薄介质, 输运建模适合. 对于常密度的高 Z 腔壁材料和靶丸的高密度低 Z 材料, 辐射自由程短, 是光性厚的, 可采用扩散建模以节省计算量. 因此, 输运与扩散耦合建模成为黑腔辐射场计算的一种很自然的选择, 对于腔内等离子体采用输运建模, 对于腔壁和靶丸采用扩散建模.

在输运扩散耦合建模的研究中, 需要解决的问题包括:

- (1) 两种物理模型耦合连接条件的具体数学表示, 其要保证连接面处的法向流连续.
- (2) 两种物理模型连接位置的确定. 要求保证计算精度的同时, 尽量减少输运区域以减少计算量.
- (3) 同一区域不同时刻不同物理模型耦合的问题. 需要解决的问题包括转换时刻的确定、物理量间的转换以及空间区域连接条件的数学表达等.

对一维定常线性输运问题, 输运与扩散的耦合已有一些研究结果, 例如针对双尺度区域分解已有严格的理论分析, 该方法可推广到二维矩形网格 (例如参见 [33-35]). 但对高维输运问题, 实用、高效且保持物理特性的扩散输运耦合建模 (计算方法) 的研究成果还很少.

3.3 Euler 和 Boltzmann 方程耦合建模与数学表达

在飞机制造、天气预报、核爆炸过程等问题数值模拟研究中, 人们往往需要了解气体在一定的区域内所形成的流场, 并以此为基础进一步研究物理机制等. 与液体、固体等不同, 大多数气体分子间间隙较大, 具有较好的可压性和流动性, 因而气体遵从自身的一系列物理模型. 在气体动力学领域中, 最常使用的是 Euler 方程组. 这一方程组建立在气体最重要的三个物理量, 密度 ρ 、运动速度 \mathbf{u} 、能量密度 E , 之上, 基于连续介质假设 (即流体所占空间可认为是连续地无间隙地充满着质点), 并通过三大守恒定律 — 质量守恒、动量守恒、能量守恒 — 推导得到, 它的具体形式如下式所示:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) &= 0, \\ \frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u} \otimes (\rho \mathbf{u})) + \nabla p &= 0, \\ \frac{\partial E}{\partial t} + \nabla \cdot (\mathbf{u}(E + p)) &= 0, \end{aligned}$$

其中 p 为气体压强, 由气体的状态关系 $p = p(\rho, e)$ 给出. 这里 e 为气体的比内能密度, 满足 $\rho e = E - \rho |\mathbf{u}|^2 / 2$. 从通常人类生活的环境至通用飞机所处的高达万米的高空, 这一物理模型均可十分精确

地描述气体流场.

然而, 当海拔高度继续增加, 进入平流层以上的中间层时 (距地球表面约 50 至 85 千米), 由于大气已经变得十分稀薄, 连续介质假设不再成立. 因而 Euler 方程组也就失去了它在常态下描述流体时所拥有的精度. 在航天飞机的设计和制造过程中, 稀薄大气的流场分布又是必须考虑的一个环节, 这就引出了气动力学的一个分支: 稀薄气体动力学. 稀薄气体动力学中的基本模型是 Boltzmann 于十九世纪六十年代所提出的一个统计物理模型, 即 Boltzmann 方程. 简而言之, Boltzmann 将气体看作许许多多相互碰撞的气体分子所构成的体系, 然后通过统计力学的手段, 将气体分子的速度分布用概率密度的形式表示出来. 方程以分布函数 $f(t; \mathbf{x}, \mathbf{v})$ 为解变量, 其中 \mathbf{x} 表示空间位置, \mathbf{v} 为微观粒子的运动速度. 方程的形式如下所示:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = Q(f, f).$$

右端项 $Q(f, f)$ 是一个二次形式的碰撞项, 用于描述粒子之间的相互作用. 显而易见的是, 除去时间变量之外, 这一方程的自变量个数是 Euler 方程组的两倍, 这导致在应用这一方程进行数值模拟时, 我们需要面对极大的计算量.

事实上, 当我们在模拟飞行器处于稀薄大气环境中周围的流体状态时, 并非在所有空间区域中都需要使用 Boltzmann 方程才能准确模拟. 在飞行器的喷气口周围, 气体仍较为稠密, 使用 Euler 方程组即可给出较为精确的模拟结果; 在远离飞行器的区域中, 气体的稀薄效应也不明显. 这些事实引导我们仅在需要使用 Boltzmann 方程的区域使用该微观模型求解, 而在绝大部分区域中使用计算量小得多的宏观模型, 即 Euler 方程组来进行计算. 这就为我们留下了三个需要解决的问题:

1. 如何确定哪一块区域需要使用计算量较大的 Boltzmann 方程进行求解;
2. 如何实现两种物理模型的耦合;
3. 当某个区域上的物理模型发生改变时, 如何实现合理的过渡.

在这三点之中, 最后一问的答案较为明显. 事实上, Euler 方程组是 Boltzmann 方程的一个简单近似, 它假设在空间的任何一点, 粒子的速度分布都服从如下形式的 Maxwell 分布:

$$f_M(\mathbf{v}) = \frac{\rho}{(2\pi RT)^{3/2}} \exp\left(-\frac{|\mathbf{v} - \mathbf{u}|^2}{2RT}\right),$$

这里 R 为气体常数, T 为气体温度. 由此可知, 当某个空间区域中所使用的模型从宏观转向微观时, 只需要用该分布函数来提供解函数即可. 反之, 当模型由微观转向宏观时, 可根据宏观与微观变量间的关系, 通过下列积分完成:

$$(\rho, \rho\mathbf{u}, E)^T = \int_{\mathbb{R}^3} (1, \mathbf{v}, |\mathbf{v}|^2/2)^T f d\mathbf{v}.$$

基于这一认识, 我们还可部分地给出第一个问题的解决方案: 当某一区域上气体的分布函数十分接近 Maxwell 分布函数时, 即可使用 Euler 方程组进行求解. 然而, 对于原本使用宏观模型求解的区域, 我们无法获知其微观分布函数的状态, 如何判断它是否需要使用微观模型求解就成为了一个难题. 目前解决这一问题的技术手段主要有两种. 一是使用 Bird 在 [36] 中所提出的“局部 Knudson 数”的概念来描述稀薄效应的强弱, 这一物理量的表达式为

$$\text{Kn}_{\text{local}} = \frac{\lambda}{Q} |\nabla Q|,$$

其中 λ 为气体分子的平均自由程, Q 是可选的某个宏观物理量, 根据问题需要可选为密度、速度或温度. 一般来说, 这个值越大, 则表示稀薄效应越强烈. 当这个数超过了一定的阈值后, 则应使用微观模

型进行求解. 文献 [37] 对这一问题进行了深入的研究, 这一方案已被 [38, 39] 等多篇文献采用. 另一方案是使用所谓“非平衡物理量”来进行稀薄效应的评定, 例如:

$$\tau_{ij} = \int (v_i - u_i)(v_j - u_j)f(\mathbf{v}) d\mathbf{v} - \rho RT\delta_{ij}, \quad q_j = \int \frac{1}{2}(v_j - u_j)|\mathbf{v} - \mathbf{u}|^2 f(\mathbf{v}) d\mathbf{v}.$$

注意这些量在 $f(\mathbf{v})$ 是 Maxwell 分布时总为零, 因而它们可作为分布函数偏离 Maxwell 分布大小的一个指标. 由于这两个表达式中仍然出现了分布函数 $f(\mathbf{v})$, 而在使用 Euler 方程组求解的区域中无法直接获得, 我们需要通过 Navier-Stokes 定律和 Fourier 定律获得它们的近似表达式, 详见 [40]. 至于具体如何使用非平衡物理量来进行稀薄效应的评估, 读者可参考文献 [41–43] 以获取详细内容.

另一个较为困难的问题是如何较好地把两种物理模型耦合在一起, 亦即如何处理两种模型的边界问题. 为了避免两种模型的交界处产生振荡, 目前较好的方式是在使用宏观模型求解的区域和使用微观模型求解的区域之间设计一块“过渡区域”. 具体而言, 我们定义一个 $[0, 1]$ 区间内取值的连续函数 $h(\mathbf{x})$, 并假设空间中的气体粒子速度分布函数具有如下形式:

$$f(t; \mathbf{x}, \mathbf{v}) = h(t; \mathbf{x})f_K(t; \mathbf{x}, \mathbf{v}) + [1 - h(t; \mathbf{x})]f_M(t; \mathbf{x}, \mathbf{v}),$$

其中 f_K 表示微观模型部分的分布函数. 不难看出, 当 $h = 0$ 时, 上式所表示的分布函数即为 Maxwell 分布, 于是模型自动退化为宏观模型, 而当 $h = 1$ 时, 该式所描述的即为微观模型. 当 h 的值介于 0 和 1 之间时, 则处于“过渡区域”之中. 函数 h 自身随时间的发展方式可由我们在上一段中所述的判断一个区域是否需要通过微观模型求解的方式给出, 而由 h 的发展方式及 Boltzmann 方程, 我们可推导出 f_K 与 f_M 的发展方程, 这样就完成了两种物理模型的多尺度耦合. 这一思路最先于 [44] 中提出, 而后在 [42, 45] 等多篇文献中都得到了应用.

4 机理与数据混合建模

随着科学研究和工程技术的不断发展, 涌现出许多不能单纯用模型或数据描述的科学问题. 对这些问题, 一方面对其原理机制的认识还不够, 相关的模型不成熟甚至没有模型; 另一方面, 相关的数据也不完备; 因此, 需要模型和数据的互补与融合进行科学计算, 以加强对这类问题的研究和认识. 混合建模的核心问题将是基于机理的模型和数据典型特征的匹配、尺度与粒度的匹配, 以及与之相关的样本数量需求分析.

4.1 PageRank 算法

PageRank 算法是一种网页排名算法, 从商业价值角度看, 也是目前最成功的可计算建模案例. PageRank 算法最早是由美国 Google 公司的创始人 Page 和 Brin 于 1998 年在美国 Stanford 大学研究新型的搜索引擎技术时所提出. 从数学的角度看, 这一算法的原始模型是惊人的简单, 然而所产生的巨大效用却不可思议. 这更凸显出数学建模在数据科学中的重要性. PageRank 算法的构造主要引自文献 [46].

PageRank 算法的基本思想是网页的重要性由所有指向它的页面的重要性之加权和所决定. 每一个指向该页面的超级链接相当于对该网页投一票, 该票数的权重由指向此网页的页面的权重所决定. 数学模型的构造则采用了网页浏览中随机点击的假设: 即上网者在进行网页浏览时以概率 q 对当前

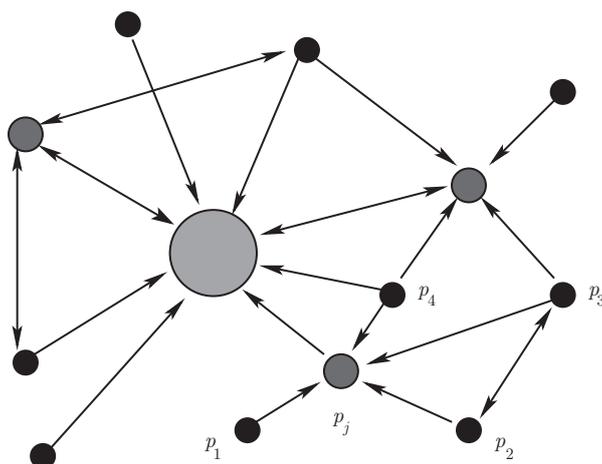


图 1 PageRank 算法进行网页排名示意图

所在网页所指向的所有页面进行等概率的点击，以概率 $1 - q$ 对整个网络所有可能页面进行等概率点击以进行下一步浏览. 这里 $0 < q < 1$ 称为阻尼系数，通常选取 q 较大，一个典型的取法是 $q = 0.85$.

以图 1 为例，假设我们考虑的网页为 p_1, p_2, \dots, p_N ，相互之间存在超级链接形成页面网络. 记第 i 个页面的出度（从第 i 个页面指向其他页面的链接数目）为 $L(p_i)$. 下面考察网页 p_j 的重要性. 所有指向 p_j 的网页有 p_1, p_2, p_3 和 p_4 . 假设网页 p_k 的重要性为 r_k . 网页 p_1 唯一的指向 p_j ，它对 r_j 的贡献是 r_1 ；网页 p_2 有两个指向外部的链接，其中之一指向 p_j ，所以它对 r_j 的贡献是 $r_2/2$ ；类似地，网页 p_3 对 r_j 的贡献是 $r_3/3$ ，网页 p_4 对 r_j 的贡献是 $r_4/3$. 根据网页浏览的随机点击假设，网页 p_j 的重要性指标由两部分组成，第一部分来源于从 p_1, p_2, p_3, p_4 到 p_j 的指向链接，具有表达式

$$q \cdot \left(r_1 + r_2 \cdot \frac{1}{2} + r_3 \cdot \frac{1}{3} + r_4 \cdot \frac{1}{3} \right),$$

第二部分来源于从网络上所有其他网页结点到 p_j 的无规随机点击，具有表达式

$$(1 - q) \sum_{k=1}^N r_k \cdot \frac{1}{N}.$$

综合上述得到

$$r_j = (1 - q) \sum_{k=1}^N r_k \cdot \frac{1}{N} + q \cdot \left(r_1 + r_2 \cdot \frac{1}{2} + r_3 \cdot \frac{1}{3} + r_4 \cdot \frac{1}{3} \right).$$

在一般情形下，定义网页的 PageRank 为向量

$$\mathbf{R} = (r_1, r_2, \dots, r_N),$$

其中 $r_i = \text{PageRank}(p_i)$ 即为该网页的 PageRank 权重. 则向量 \mathbf{R} 满足

$$\mathbf{R} = \mathbf{R} \cdot \mathbf{P},$$

这里 P 矩阵为

$$P = (1 - q) \cdot \frac{1}{N} I_{N \times N} + q \cdot \begin{pmatrix} l_{11} & l_{12} & \dots & l_{1N} \\ l_{21} & l_{22} & \dots & l_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ l_{N1} & l_{N2} & \dots & l_{NN} \end{pmatrix},$$

其中如果 $L(p_i) \neq 0$ 则定义 $l_{ij} = 1/L(p_i)$, 这里指标 j 仅取页面 p_i 所指向的页面; 否则取 $l_{ij} = 0$. 注意到 P 矩阵是一个本原随机矩阵, 即 P 为非负矩阵, 行和为 1, 且 P 矩阵元素均非零, 而向量 \mathbf{R} 实际上是该矩阵的特征值为 1 (模最大特征值) 的左特征向量, 也就是该随机矩阵所对应的 Markov 链的不变分布. 根据 Perron-Frobenius 定理, \mathbf{R} 必存在唯一. 为计算网页排名向量 \mathbf{R} , 根据上述解释, 一个简单的方法是采用线性代数中计算模最大特征值所对应的特征向量的幂法. 即设定初始迭代向量 \mathbf{R}_0 , 通常取为 $\mathbf{R}_0 = (1/N, 1/N, \dots, 1/N)$, 然后根据

$$\mathbf{R}_{n+1} = \mathbf{R}_n \cdot P$$

迭代直到收敛即可.

4.2 功能磁共振成像

人脑是人体最重要的器官之一, 对于人脑功能的探求无疑是非常有意义的事情. 长久以来, 科学家们就注意到这样的事实: 即人脑的功能反映在大脑皮层是按空间分区的, 在脑内次级结构也是按空间分隔的. 研究脑功能映射有许多成功的模式, 在众多的模式中最成功的当属功能磁共振成像 (fMRI) [47, 48]. fMRI 了解人类大脑的工作机制、认知过程、情绪过程等问题都有重要意义. fMRI 是目前研究临床脑疾病最重要的技术之一, 利用 fMRI 技术, 临床研究人员已经能够获取脑激活、功能连接、网络属性、能量代谢、生化反应等多方面的信息, 为全面深入了解包括精神分裂症、老年痴呆症、儿童注意缺陷多动障碍、癫痫、中风、康复、药物成瘾等一系列临床问题的神经生理病理学机制提供了重要帮助.

fMRI 技术是多学科交叉合作的产物, 其中成像原理涉及仪器仪表、电子工程、物理学等学科, 分析仪器测量的结果并形成图象主要涉及计算科学和统计学. 目前应用最广泛、成就最显著的脑功能成像技术称为基于血氧水平 (BOLD) 的功能磁共振成像 (fMRI), 简记为 BOLD-fMRI. 其原理基于 “BOLD 效应” [49], 大意为: 磁共振扫描仪可以直接或间接得到人脑血氧水平变化的信号. 人脑活动不同, 血氧代谢水平就不同. 当特定脑区被激活时, 相应的局部脑组织血流和代谢增加, BOLD 信号相对增强. 这些信号就反映了相应脑神经细胞活动的变化, 由此可测量出兴奋区域, 这个兴奋区域就与特定的脑功能有关, 再通过 BOLD 信号的空间编码形成图象.

4.2.1 基本原理

普通临床用的 MRI 信号几乎都来自组织液中的质子, 图像强度主要取决于质子的密度, 但是水分子周围局部环境也对它有很大的影响. 质子受到一个射频磁场脉冲的激发后, 它的磁化方向不再与 MRI 磁体的静态磁场方向一致, 需要较长的时间 (大约从零点几秒到几秒) 才能回到原来的方向. 在这段时间里, 和静态磁场垂直的磁化分量在被扫描的物体周围的导线中产生一个感应信号电压. 如果水分子的质子在完全恢复之前再次被激励, 则产生一个相对较小的信号. 恢复率称为纵向弛豫时间 T_1 , 不同组织中质子的 T_1 不同.

改变射频脉冲的重复时间 (TR), T_1 较长和较短的组织间对比会发生显著变化. 为了观察 MRI 信号, 质子磁化方向必须偏离主磁场方向, 在横断面上生成一个沿轴进动的磁化分量. 为使这个信号最大, 横断面上的磁化向量产生的相角在围绕物体方向保持恒定, 这样才能将每个质子的磁化分量叠加起来. 然而, 每个质子自旋的磁环境不同使它们以稍微不同的频率进动, 使相角分离, 从而使信号随时间减小. 信号基本上按指数规律衰减的, 衰减的速度由时间常数 T_2 (横向弛豫时间) 决定.

横向净磁场的衰减总是比纵向净磁场的衰减要快. 而且, 由于体内顺磁粒子 (如某些 MRI 对比剂) 的存在或由于物体本身的空间不均匀性引起物体周围的磁场变化都会使相角进一步分散, 使信号更快衰减. 这个附加的弛豫时间定义为 \hat{T}_2 . 总的信号衰减由弛豫时间常数 T_2^* 所决定, 它们之间的关系是:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{\hat{T}_2}.$$

功能磁共振就是利用磁场不均匀性对衰减信号进行测量. 因为横向净磁场的衰减非常快, 所以可以在非常短的时间内检测到信号, 这就提供了很好的时间分辨率. 通常使用回波技术对衰减信号进行测量, 自旋回波技术用于测量 T_2 信号, 梯度回波技术用于测量 T_2^* 信号.

磁共振现象的数学模型可以用下列 Bloch-Torrey 方程描述:

$$\begin{aligned}\frac{\partial M_x}{\partial t} &= \gamma(\mathbf{M} \times \mathbf{B})_x - \frac{M_x}{T_2} + \nabla \cdot (\mathbf{D} \nabla M_x), \\ \frac{\partial M_y}{\partial t} &= \gamma(\mathbf{M} \times \mathbf{B})_y - \frac{M_y}{T_2} + \nabla \cdot (\mathbf{D} \nabla M_y), \\ \frac{\partial M_z}{\partial t} &= \gamma(\mathbf{M} \times \mathbf{B})_z + \frac{M_0 - M_z}{T_1} + \nabla \cdot (\mathbf{D} \nabla M_z),\end{aligned}$$

其中, M_x , M_y 和 M_z 是磁化强度矢量 \mathbf{M} 的三个分量, \mathbf{B} 是外加磁场强度矢量, γ 是原子磁旋比, T_1 是纵向弛豫时间, T_2 是横向弛豫时间, \mathbf{D} 是二阶扩散张量, 即一个 3×3 的对称矩阵 (d_{ij}). 磁共振成像的目标, 就是利用测量被测样本发射出的电磁波信号反求 Bloch-Torrey 方程中的各个参数, 如 T_1 , T_2 等.

4.2.2 数据处理

在 fMRI 信号的数据处理中, 涉及到时空数据的许多统计推断问题. 人脑是三维的, 设要研究的某个局部脑区为 Ω , 把 Ω 划分为 N 个单元, 称为体素, 记为 $d \in \Omega$. 在磁共振仪器的扫描过程中, 记体素 d 在 t_1, \dots, t_n 时间的信号值为 $Y_i(d)$ ($i = 1, \dots, n$, 至少取数百个). 对于固定的 d , $Y_i(d)$ 为一时间序列, 而 $d \in \Omega$ 可以是任一设定的体素. 因此数据集 $Y_i(d)$, $i = 1, \dots, n; d \in \Omega$ 与时间空间都有关, 是一种特定的高维时空数据. 而且噪声来源比较多, 如心跳、呼吸、血管收缩、周围环境等等. 文献中有许许多多关于这类“成像数据”的处理方法, 其中有代表性的一本著作是《统计参数图脑功能成像分析》(见 [50]), 该书详细阐述了脑信号分析的概念和过程, 全面系统地介绍了各种脑成像数据的分析和处理方法. 这本书所介绍的数据处理方法就是 SPM 软件的主要依据, SPM 软件主要基于两种方法: 第一种是应用一般线性模型拟合数据, 并通过假设检验及减压的 p -值测定脑激活区域; 第二种通过应用随机场理论处理各体素之间的多重比较问题, 整合扫描过程中得到的各个激活区域.

最近, 美国加州大学 Berkeley 分校科学家应用机器学习的方法, 通过 fMRI 信号的稀疏线性模型 [51], 并融合人视觉对图象的空间、定向、频率和对比度的非线性调谐机制, 成功地实现了“现代读心术”: 从被试者的 fMRI 成像, 可以“看到”被试者大脑中“正在思考”的图象.

功能磁共振成像的物理过程认识不够, 数据也不完备, 需要模型和数据的互补与融合进行科学计算并成像, 现在已经有许多相关软件, 已经成为临床脑疾病最重要的技术.

5 结束语

本文着重讨论和分析了可计算建模的内涵, 并以典型例子说明了可计算建模在科学与工程计算中的重要作用. 当前, 许多国家重大需求和重要科学前沿研究中的复杂科学计算问题都不是单一学科的问题, 解决这些复杂问题一方面需要相关领域专家和数学专家合作, 开展基于相关领域知识和数学方法的计算建模研究, 另一方面也需要数学专家和相关领域专家合作, 针对模型的背景与特点, 研究高效实用的算法和可计算模型所需解决的关键数学问题, 只有这样才能有效解决实际的科学计算问题.

因此, 在科学与工程计算中, 应以可计算建模为纽带, 将算法研究与解决实际问题紧密结合起来, 积极促进两者的融合与发展、数学与其它学科的交叉, 促进数学研究人员与其它领域的专业人员密切合作, 共同进行可计算建模与共性算法的研究, 造就一批高水平的科学与工程计算复合型人才, 推动科学与工程计算乃至科学技术的跨越发展.

致谢 感谢参与组织自然科学基金委员会重大研究计划“高性能科学计算基础算法与可计算建模”的各位专家, 特别是崔俊芝、郭柏灵院士, 陈志明、程晋、汤涛教授, 雷天刚、何成、赵桂萍研究员等. 感谢蔡振宁同学、蒋凯博士、李敬宏研究员、李铁军教授、林霖博士、明平兵研究员、杨志坚教授等提供可计算建模的相关例子.

参考文献

- 1 Hohenberg P, Kohn W. Inhomogeneous electron gas. *Phys Rev B*, 1964, 136: 864–871
- 2 Kohn W, Sham L. Self-consistent equations including exchange and correlation effects. *Phys Rev A*, 1965, 140: 1133–1138
- 3 Ceperley D M, Alder B J. Ground state of the electron gas by a stochastic method. *Phys Rev Lett*, 1980, 45: 566–569
- 4 Perdew J P, Zunger A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys Rev B*, 1981, 23: 5048–5079
- 5 Becke A D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A*, 1988, 38: 3098–3100
- 6 Lee C, Yang W, Parr R G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B*, 1988, 37: 785–789
- 7 Perdew J P, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Phys Rev Lett*, 1996, 77: 3865–3868
- 8 Becke A D. A new mixing of Hartree Fock and local density functional theories. *J Chem Phys*, 1993, 98: 1372–1377
- 9 Hamann D R, Schlüter M, Chiang C. Norm-conserving pseudopotentials. *Phys Rev Lett*, 1979, 43: 1494–1497
- 10 Troullier N, Martins J L. Efficient pseudopotentials for plane-wave calculations. *Phys Rev B*, 1991, 43: 1993–2006
- 11 Vanderbilt D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys Rev B*, 1990, 41: 7892–7895
- 12 Goedecker S. Linear scaling electronic structure methods. *Rev Mod Phys*, 1999, 71: 1085–1123
- 13 Kohn W. Density functional and density matrix method scaling linearly with the number of atoms. *Phys Rev Lett*, 1996, 76: 3168–3171
- 14 van der Waals J D. Over de Continuïteit van dpe Gas- en Vloeistoftoestand. Thesis, Leiden, 1873
- 15 Landau L D, Lifshitz E M. *Statistical Physics*. Reading, Mass: Addison-Wesley, 1969
- 16 Stanley H E. *Introduction to Phase Transition and Critical Phenomena*. New York: Oxford University Press, 1987
- 17 Miller P D. *Applied Asymptotic Analysis*. Providence, RI: Amer Math Soc, 2006
- 18 Jiang K, Wang C, Huang Y, Zhang P. Discovery of New Metastable Patterns in Diblock Copolymers. *Commun Comput Phys*, submitted
- 19 Bird R B, Curtiss C F, Armstrong R C, et al. *Dynamics of Polymeric Liquids, Vol. 2: Kinetic Theory*, 2nd ed. New York: Wiley, 1987

- 20 Ilg P, Karlin I V, Öttinger H C. Canonical distribution functions in polymer dynamics, I: Dilute solutions of flexible polymers. *Physica A*, 2002, 315: 367–385
- 21 Wang H, Li K, Zhang P. Crucial properties of the moment closure model FENE-QE. *J Non-Newtonian Fluid Mech*, 2008, 150: 80–92
- 22 E W. *Principles of Multiscale Modeling*. New York: Cambridge Univ Press, 2011
- 23 Tadmor E B, Ortiz M, Phillips R. Quasicontinuum analysis of defects in solids. *Phil Mag A*, 1996, 73: 1529–1563
- 24 Born M, Huang K. *Dynamical Theory of Crystal Lattices*. New York: Oxford University Press, 1954
- 25 Shenoy V B, Miller R, Tadmor E B, et al. An adaptive finite element approach to atomic scale mechanics—the quasicontinuum method. *J Mech Phys Solids*, 1999, 47: 611–642
- 26 Ming P, Yang J Z. Analysis of a one-dimensional nonlocal quasicontinuum method. *Multiscale Model Simul*, 2009, 7: 1838–1875
- 27 W E, Lu J, Yang J Z. Uniform accuracy of the quasicontinuum method. *Phys Rev B*, 2006, 74: 214115
- 28 Shimokawa T, Mortensen J J, Schiøz J, et al. Matching conditions in the quasicontinuum method: removal of the error introduced at the interface between the coarse-grained and fully atomistic region. *Phys Rev B*, 2004, 69: 214104
- 29 Miller R E, Tadmor E B. A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods. *Modelling Simul Mater Sci Eng*, 2009, 17: 053001
- 30 Li X, W E, Yang J Z. A multiscale coupling method for the modeling of dynamics of solids with application to brittle cracks. *J Comput Phys*, 2010, 229: 3970–3987
- 31 Mihalas D, Mihalas B W. *Foundations of Radiation Hydrodynamics*. New York: Oxford Univ Press, 1984
- 32 Pomraning G C. *The Equations of Radiation Hydrodynamics*. New York: Pergamon Press, 1973
- 33 Golse F, Levermore C D, Shi J. A domain decomposition analysis for a two-scale linear transport problem. *Math Model Num Anal*, 2003, 37: 869–892
- 34 Jin S, Yang X, Yuan G. A domain decomposition method for a two-scale transport equation with energy flux conserved at the interface. *Kinet Relat Models*, 2008, 1: 65–84
- 35 Pomraning G C, Foglesong G M. Transport-diffusion interfaces in radiative transfer, *J Comp Phys*, 1979, 32: 420–436
- 36 Bird G A. *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*. Oxford: Clarendon Press, 1994
- 37 Wang W, Boyd I D. Predicting continuum breakdown in hypersonic viscous flows. *Phys Fluids*, 2003, 15: 91–100
- 38 Schwartzentruber T E, Boyd I D. A hybrid particle-continuum method applied to shock waves. *J Comput Phys*, 2006, 215: 402–416
- 39 Wu J, Lian Y, Cheng G, et al. Development and verification of a coupled DSMC-NS scheme using unstructured mesh. *J Comput Phys*, 2006, 219: 579–607
- 40 Struchtrup H. *Macroscopic Transport Equations for Rarefied Gas Flows: Approximation Methods in Kinetic Theory*. New York: Springer, 2005
- 41 Boyd I D, Chen G, Candler G V. Predicting failure of the continuum fluid equations in transitional hypersonic flows. *Phys Fluids*, 1995, 7: 210–219
- 42 Degond P, Dimarco G, Mieussens L. Analysis of an asymptotic preserving scheme for linear kinetic equations in the diffusion limit. *SIAM J Numer Anal*, 2010, 48: 1474–1491
- 43 Garcia A L, Bell J B, Crutchfield W Y, Alder B J. Adaptive mesh and algorithm refinement using direct simulation Monte Carlo. *J Comput Phys*, 1999, 154: 134–155
- 44 Degond P, Jin S. A smooth transition model between kinetic and diffusion equations. *SIAM J Numer Anal*, 2004, 42: 2671–2687
- 45 Cai Z, Li R. An h -adaptive mesh method for Boltzmann-BGK/hydrodynamics coupling. *J Comput Phys*, 2010, 229: 1661–1680
- 46 Brin S, Page L. The anatomy of a large scale hypertextual web search engine. *Comput Netw ISDN Syst*, 1998, 30: 107–117
- 47 Goldstein S J, Wekstein D R, Kirkpatrick C. Cerebral computed tomography and magnetic-resonance imaging in the functional centenarian. *Amer J Neuroradiology*, 1985, 6: 467–467
- 48 Flannigan B D, Bradley W G, Maziotta J C, et al. Magnetic-resonance imaging of the brain-stem-normal structure and basic functional anatomy. *Radiology*, 1985, 154: 375–383
- 49 Ogawa S, Menon R S, Tank D W, et al. Functional brain mapping by blood oxygenation level-dependent contrast magnetic-resonance-imaging—A comparison of signal characteristics with a biophysical model. *Biophysical J*, 1993, 64: 803–812
- 50 Friston K, Ashburner J, Kiebel S. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. London: Academic Press, 2007

- 51 Vu V Q, Ravikumar P, Naselaris T, et al. Encoding and decoding V1 fMRI responses to natural images with sparse nonparametric models. *Ann Appl Statist*, 2011, 5: 1159–1182

Computable modeling

Ji PeiWen, JIANG Song & ZHANG PingWen

Abstract Computable modeling is to construct or simplify models using the related information through the precision of problems, the goal of which is to reduce the computational cost and increase the efficiency, such that the model could be computable using the computer nowadays. The computable modeling is an important component of scientific computing. This article is to introduce the concept of computable modeling by some examples.

Keywords scientific computing, computable modeling, basic algorithm

MSC(2010) 65Z05, 65Z99

doi: 10.1360/012012-144