布尔巴基研讨会 2022年11月
75e 年，2022–2023年，第 1199期

arXiv:2301.01511v1 [math.DS] 4 Jan 2023

逐点遍历理论：例子与熵[根据 Jean Bourgain 的讲座改编]

作者：**Ben Krause**

概述

作为离散调和分析研究动机的逐点遍历理论，其根基是伯克霍夫的经典定理 [2], ，可描述如下：

> 对于概率空间 (X, μ) 上任意遍历（即“具有足够随机化作用”）的保测变换 τ，以及任意可积函数$f \in L^1(X, \mu)$，几乎处处（ μ），都可以通过考察切萨罗和来恢复 f 的平均值

$$\frac{1}{N}\sum_{n\leq N} f(\tau^n x) \to \int_X f \, d\mu \ \mu - \text{a.e.}$$

非正式地说，该定理表明，可以通过考察 f 的空间平均值，

$$\int_X f \, d\mu,$$

考察时间平均

$$\Big\{\frac{1}{N}\sum_{n\leq N} f(\tau^n x)\Big\},$$

它们是通过在“时刻” $\{\tau^n x\}$ 对函数 f 进行“采样”并取适当的平均值(1)而得到的。

逐点遍历理论中的一个经典问题，涉及平均值极限行为几乎处处存在的问题

$$\frac{1}{N}\sum_{n=1}^{N} \tau^{a_n} f \tag{1}$$

其中 $\{a_n\}$ 是“稀疏”的；按照惯例，此处及下文中我们用 $\tau^k f$ 表示函数

$$x \mapsto f(\tau^k x).$$

(1)~~即使在 τ不~~是遍历的情况下，时间平均 $\{\frac{1}{N}\sum_{n\leq N}\tau^n f(x)\}$仍然几乎处处收敛 μ。

当序列 $\{a_n\}$ 的下密度有正下界时

$$\liminf \frac{|\{n : a_n \leq N\}|}{N} > 0,$$

收敛很容易展示出来，而经典问题则是关于是否存在密度为零的序列 $\{a_n$,

$$\lim \frac{|\{n : a_n \leq N\}|}{N} = 0,$$

使得其平均值（1）几乎处处收敛。在 [2], 中，构造了这样一个序列：它由一段段较长的自然数区块组成，区块之间隔着长得多的间隙，接着是稍长一些的区块，再接着是更长的间隙，依此类推。尤其是，这个序列的上 Banach 密度为 1

$$d^*(\{a_n\}) := \limsup_{|I|\to\infty \text{ an interval}} \frac{|\{a_n\} \cap I|}{|I|} = 1.$$

然而，问题仍然是：是否存在上 Banach 密度为零的序列，$\{a_n\}$ 使得 $d^*(\{a_n\}) = 0$ 的平均值（1）能够证明几乎处处收敛。特别是，Furstenberg 最先明确提出的经典问题， [16], 另见 [1], 是沿平方数取平均值是否

$$\frac{1}{N}\sum_{n=1}^{N} \tau^{n^2} f$$

几乎处处逐点收敛，最初是对 $f \in L^2(X$而言。在突破性工作中，[5]； [6]； [9], Bourgain 对这个问题给出了肯定回答，并证明了对于任意多项式序列，（1）几乎处处收敛,

$$\{a_n = P(n)\}, \quad P \in \mathbb{Z}[\cdot],$$

以及任意的 $f \in L^p(X)$ $p > 1$，对于任意 σ-有限测度空间 X；这一结果后来被证明是尖锐的 [11]； [25]。

定理 0.1。——设(X, μ) 是一个 σ-有限测度空间， $\tau : X \to X$ 是一个保测变换，而 $P \in \mathbb{Z}[\cdot]$ 是一个系数为整数的多项式。那么对每个 $1 < p < \infty$

$$\frac{1}{N}\sum_{n=1}^{N} \tau^{P(n)} f$$

几乎处处收敛（相对于 μ）。

尽管逐点收敛问题属于定性问题，Bourgain 的洞见在于量化收敛发生的速率——随后运用 Calderón 首创的一种抽象转移论证 [12] ，从一个“通用”的保测系统中推导出这些定量估计。通过考察形如

$$\mathbb{Z} \ni n \mapsto \tau^n f(x), \quad x \in X \text{ fixed}$$

的序列，并利用 τ 的保测性质，Bourgain 得以将问题归结为在配备计数测度和移位（$\mathbb{Z}$, $|\cdot|$, τ: $x \mapsto x-1$）的整数情形下证明这些估计。

尤其是，Bourgain 关注平均算子的振荡的定量估计

$$\frac{1}{N}\sum_{n=1}^{N} f(x-P(n)), \tag{2}$$

最初应用于 $\ell^2(\mathbb{Z})$-函数。理解 (2) 的一个自然视角，是将其看作 f与

$$K_N(x) := \frac{1}{N}\sum_{n=1}^{N} \delta_{P(n)}(x)$$

其中 δ_m 表示位于 $m \in \mathbb{Z}$处的点质量；由于这个问题基于 $\ell^2(\mathbb{Z})$，因此自然会采用傅里叶变换方法，而分析的关键在于理解这些指数和

$$\frac{1}{N}\sum_{n\le N} e^{-2\pi i\beta\cdot P(n)},$$

这可通过解析数论中的圆法 实现；逐点收敛这一“软性”分析问题，与通过解析数论方法在整数/欧几里得空间上获得“硬性”分析估计之间的相互作用，是逐点遍历理论和离散调和分析这两个领域的典型特征。

我最初是通过阅读 [36], 才理解 Bourgain 的工作；我认为它对定理 0.1 作了非常精彩的阐释；这些笔记旨在补充 [36] ，尝试解释 Bourgain 论证背后的动机。

因此，为了便于说明，我们将把重点从证明定理 0.1稍微转移开来，转而关注相关的极大估计，具体考虑 $L^2(X)$.

定理 0.2。——设 (X,μ) 是一个 σ-有限测度空间， τ： $X \to X$ 是一个保测变换，并且 $P \in \mathbb{Z}[\cdot]$ 是一个系数为整数的多项式。那么存在一个绝对常数$\mathbf{C}$，与(X,μ,τ)无关，使得

$$\| \sup_N \left| \frac{1}{N}\sum_{n=1}^{N} \tau^{P(n)} f \right| \|_{L^2(X)} \le \mathbf{C}\cdot \|f\|_{L^2(X)}.$$

根据 Calderón 的迁移原理，定理 0.2 可由整数上的相应估计推出：若定义

$$\mathscr{M} f(x) := \sup_N \left| \frac{1}{N}\sum_{n=1}^{N} f(x-P(n)) \right|, \tag{3}$$

接下来，我们将重点转向证明以下估计

定理 0.3. — 对于任意 $P \in \mathbb{Z}[\cdot]$, ，以下范数不等式成立：存在一个绝对常数$\mathbf{C}$ 使得

$$\|\mathscr{M} f\|_{\ell^2(\mathbb{Z})} \leq \mathbf{C} \cdot \|f\|_{\ell^2(\mathbb{Z})}.$$

下面，按照 [36], 的思路，我们将讨论范围限定在

$$P(n) = n^d,$$

因为这样既能避免一些数论技术细节，又能抓住问题的本质。

0.0.1. 记号。——本文中，我们将复指数$e(t)$简记为$= e^{2\pi i t}$，因此可以分别表示欧几里得空间和整数上的傅里叶变换为

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) \cdot e(-\xi x)\ dx, \quad g^{\vee}(x) = \int_{\mathbb{R}} g(\xi) \cdot e(\xi x)\ d\xi$$
$$\hat{f}(\beta) = \sum_{n} f(n) \cdot e(-\beta n), \quad g^{\vee}(n) = \int_{\mathbb{T}} g(\beta) \cdot e(\beta n)\ d\beta.$$

我们令

$$\phi_k(t) := 2^{-k} \cdot \phi(2^{-k} \cdot t)$$

表示通常的 L^1-归一化二进缩放；对于频率 θ，令

$$\mathrm{Mod}_\theta g(x) := e(\theta x) \cdot g(x) \tag{4}$$

因此

$$\widehat{\mathrm{Mod}_\theta g}(\beta) = \hat{g}(\beta - \theta),$$

并回顾 Hardy–Littlewood 极大算子

$$M_{\mathrm{HL}} f(x) := \sup_{r>0} \frac{1}{2r} \int_{-r}^{r} |f(x-t)|\ dt \quad \text{or} \quad := \sup_{N \geq 0} \frac{1}{2N+1} \sum_{n=-N}^{N} |f(x-n)|;$$

虽然我们用相同的记号指代连续和离散极大算子，但具体使用哪种形式会由上下文明确。

我们令 $[N]$:=**{1**，**.....**， N**}**，并简记 $\sum_{n \leq N}$:= $\sum_{n=1}^{N}$。我们用符号 $\mathbf{c}$ 表示适当小的常数，其值始终与零保持一定距离；用 $\mathbf{C}$ 表示适当大的常数，其值始终有上界。如果需要这些常数依赖于某些参数，我们就用下标表示，因此 $\mathbf{c}_d$ 是一个依赖于 d 的小常数。我们用 $X = O(Y)$ 表示 $|X| \leq \mathbf{C} \cdot Y$，并类似地定义 $X = O_d(Y)$。

最后，我们将使用启发式记号

$$f \ \text{“}=\text{”}\ g$$

表示实质等价：在可接受的误差范围内， f 和 g 表现出相同类型的行为。

1. 离散情形的复杂性

在开始讨论定理 0.3之前，我们先来解释一下，为什么我们认为这可能是一个颇具挑战性的问题。

对于带有“线性”特征的问题，离散理论基本上与连续理论相呼应

$$\sup_r \frac{1}{r}\int_0^r |f(x-t)|\,dt \quad \text{“}=\text{”} \quad \sup_N \frac{1}{N}\sum_{n=1}^{N} |f(x-n)|$$

这可以通过考察形如 $F(\lfloor x \rfloor)$的函数，并利用实变量极大函数的伸缩不变性，将注意力缩小到在单位尺度上为常数的实变量函数来验证。

一旦线性结构被破坏，问题就会变得复杂得多。在这种情况下，我们考虑一个简单的例子：沿曲线 $t \mapsto t^d$的Hardy–Littlewood极大函数。连续极大函数

$$Mf := M_d f := \sup_r \Big|\frac{1}{r}\int_0^r f(x-t^d)\,dt\Big| = \sup_r \Big|\frac{1}{r}\int_0^{r^d} f(x-t)\,\frac{1}{dt^{1-1/d}}\,dt\Big|, \tag{5}$$

只是通过逐点主控不等式得到的 M_{HL} 的一个加权版本

$$\begin{aligned}\frac{1}{r}\int_0^{r^d} |f(x-t)|\frac{1}{dt^{1-1/d}}\,dt &\le \sum_{j=1}^{\infty} 2^{-j/d}\cdot\Big(\frac{2^{j/d}}{r}\int_{2^{-j}\cdot r^d}^{2^{1-j}\cdot r^d} |f(x-t)|\frac{1}{dt^{1-1/d}}\,dt\Big)\\ &\le \mathbf{C}/d\cdot\sum_{j=1}^{\infty} 2^{-j/d}\cdot\Big(\frac{2^j}{r^d}\int_{2^{-j}\cdot r^d}^{2^{1-j}\cdot r^d} |f(x-t)|\,dt\Big) \le \mathbf{C}/d\cdot\sum_{j=1}^{\infty} 2^{-j/d}\cdot M_{\mathrm{HL}} f(x)\\ &\le \mathbf{C}\cdot M_{\mathrm{HL}} f(x)\end{aligned} \tag{6}$$

另一方面，在研究.....时并没有这种技巧可用

$$\mathscr{M} f(x) := \mathscr{M}_d f(x) := \sup_N \Big|\frac{1}{N}\sum_{n\le N} f(x-n^d)\Big|,$$

由于存在最小尺度——在离散情形中，变量的无穷小变化并无真正的对应物。

转到傅里叶侧实际上凸显了这一差异。利用非负性后，我们可以将 M和 $\mathscr{M}$ 都表示为对一个傅里叶乘子稀疏序列取极大值的极大算子。我们先从 M开始：

$$Mf(x) := \sup_k \Big|\Big(V_k(\xi)\hat{f}(\xi)\Big)^{\vee}(x)\Big|,$$

其中

$$V_k(\xi) := \int_0^1 e(-\xi 2^{dk} t^d)\,dt, \tag{7}$$

使得

$$V_k(\xi) = \int_0^1 e(-\xi 2^{dk}t)\frac{1}{dt^{1-1/d}}\,dt =: \widehat{\mu}(2^{dk}\xi) = \begin{cases} 1 + O\big(2^{dk}|\xi|\big) \\ O\Big(\big(2^{dk}|\xi|\big)^{-1/d}\Big), \end{cases} \tag{8}$$

这可通过在原点附近对指数函数进行泰勒展开，并对第二个估计应用驻相原理（巧妙地进行分部积分）得出。上文中，我们令

$$\mu(t) := \mu_d(t) := \frac{1}{dt^{1-1/d}} \cdot \mathbf{1}_{(0,1]}. \tag{9}$$

这一分析表明，乘子 V_k 会竭力表现得像 $\widehat{\varphi_{dk}}$；例如，对于满足 $\widehat{\varphi}(0) = 1$ 的施瓦茨函数 $\varphi \geq 0$，本例中有类似的估计：

$$\widehat{\varphi_{dk}}(\xi) = \begin{cases} 1 + O\big(2^{dk}|\xi|\big) \\ O\Big(\big(2^{dk}|\xi|\big)^{-100}\Big) \end{cases} \tag{10}$$

（例如）；与（7）比较。现在，用更强的 ℓ_k^∞-范数替换 $\{(\mu_{dk} - \varphi_{dk}) * f)\}$，我们得到 ℓ_k^2-范数

$$\begin{aligned} Mf &\leq \sup_k |\varphi_{dk} * f| + \sup_k |(\mu_{dk} - \varphi_{dk}) * f| \\ &\leq \mathbf{C} \cdot M_{\mathrm{HL}} f + \Big(\sum_k |(\mu_{dk} - \varphi_{dk}) * f|^2\Big)^{1/2} \\ &=: \mathbf{C} \cdot M_{\mathrm{HL}} f + Sf, \end{aligned} \tag{11}$$

其中 Sf 是一种所谓的平方函数，专门用于研究基于 L^2 的问题。事实上，我们利用帕塞瓦尔定理来估计

$$\begin{aligned} \|Sf\|_{L^2(\mathbb{R})}^2 &= \|\Big(\sum_k |(\mu_{dk} - \varphi_{dk}) * f|^2\Big)^{1/2}\|_{L^2(\mathbb{R})}^2 = \sum_k \|(\mu_{dk} - \varphi_{dk}) * f\|_{L^2(\mathbb{R})}^2 \\ &= \sum_k \|(V_k - \widehat{\varphi_{dk}}) \cdot \hat{f}\|_{L^2(\mathbb{R})}^2 = \int \sum_k |V_k(\xi) - \widehat{\varphi_{dk}}(\xi)|^2 \cdot |\hat{f}(\xi)|^2\,d\xi \\ &\leq \sup_\xi \sum_k |V_k(\xi) - \widehat{\varphi_{dk}}(\xi)|^2 \cdot \|\hat{f}\|_{L^2(\mathbb{R})}^2 \\ &\leq \mathbf{C} \cdot \sup_\xi \sum_k \min\{2^{kd}|\xi|, (2^{kd}|\xi|)^{-1/d}\}^2 \cdot \|f\|_{L^2(\mathbb{R})}^2 \\ &\leq \mathbf{C}_d \cdot \|f\|_{L^2(\mathbb{R})}^2, \end{aligned} \tag{12}$$

利用 $\widehat{\varphi_{dk}}$(ξ满足与 V_k 相同估计这一事实，即（7），因此对于$|\xi| \leq \mathbf{C}\cdot\ 2^{-dk}$

$$V_k(\xi) - \widehat{\varphi_{dk}}(\xi) = \Big(1 + O(2^{dk}|\xi|)\Big) - \Big(1 + O(2^{dk}|\xi|)\Big) = O(2^{dk}|\xi|)$$

并且当 $|\xi| > \mathbf{C}\cdot\ 2^{-dk}$

$$V_k(\xi),\ \widehat{\varphi_{dk}}(\xi) = O((2^{dk}|\xi|)^{-1/d}).$$

如果我们对离散算子 $\mathscr{M}$ 尝试同样的技巧，

$$\mathscr{M} f(x) = \sup_k |K_k * f(x)|$$

其中

$$K_k(x) := \frac{1}{2^k} \sum_{n \leq 2^k} \delta_{n^d}(x), \tag{13}$$

我们也可以将 $\mathscr{M}$ 表示为一个极大乘子算子

$$\mathscr{M} f(x) = \sup_{k \geq 0} |\Big(\widehat{K_k}(\beta)\hat{f}(\beta)\Big)^{\vee}(x)|,$$

其中乘子 K_k 的形式与 $\{V_k\}$不同：

$$\Big\{\widehat{K_k}(\beta) := \frac{1}{2^k} \sum_{m \leq 2^k} e(-\beta m^d)\Big\}_{k \geq 0}.$$

每个乘子都是一个韦伊和，分析它需要用到 Hardy 和 Littlewood 所谓的圆法。正如我们下面将看到的，每个乘子

$$\widehat{K_k}(\beta)$$

当 β 与某个有理数“相差不超过k”，且该有理数的分母“较k小”时，也就是说，当 β 位于所谓的“k主弧”上时，会很大且具有重要意义；而当 β 位于其补集“k次弧”上时，则“k可忽略”。特别是，在分析相关乘子时，我们会遇到微妙的算术问题；这与欧几里得情形形成对比，在后者中，我们可以完全根据频率变量的大小 来理解乘子。换言之，欧几里得情形中的分析完全取决于频率变量到特定零频率的距离；而为了理解乘子 K_k（β），我们会遇到多频率问题。实质上，对其进行估计的主要工作

$$\|\mathscr{M} f\|_{\ell^2(\mathbb{Z})} \leq \mathbf{C} \cdot \|f\|_{\ell^2(\mathbb{Z})},$$

归根结底，就是要克服这些多频率带来的复杂问题。

2. 例子

在下文中，我们可以并且将假设 k 足够大，具体取决于 d。为了理解

$$\mathscr{M} f := \sup_k |K_k * f|,$$

我们先通过研究一些例子来建立直觉：

实数轴的伸缩不变性使我们能够研究5或 M_{HL}，并使用尺度为单位尺度的例子；而 $\mathbb{Z}$上则不存在这样的伸缩不变性。相反，与“放大观察”粗略类似的做法是转而考察一个算术级数。当然，这种类比并不精确，因为算术级数同时由间隔大小

和直径决定。因此，我们首先分析13作用于函数时的行为

$$\varphi_{Q,N} := \mathbf{1}_{Q\mathbb{Z}} \cdot \varphi(\cdot/N) \tag{14}$$

其中 φ 是一个光滑凸起函数，我们将 $Q \leq N^{1/2}$ 视为；注意这个近似

$$\|\varphi_{Q,N}\|_{\ell^2(\mathbb{Z})} \approx (N/Q)^{1/2}. \tag{15}$$

将问题转到算术级数时，一个常见的简化假设是令间隔大小为素数，因为这样可以消除各种算术技术细节，所以下文我们也将如此处理。

考虑到这些化简，我们开始进行计算。

2.1. 示例

出于技术上的原因，我们将用完整卷积算子 K_k的光滑“上半部分”来代替它；也就是说，对于光滑的 $\mathbf{1}_{[1,2]} \leq \phi \leq \mathbf{1}_{[1/2,4]}$，我们考虑

$$K_k' := \sum_n \phi_k(n) \cdot \delta_{n^d}. \tag{16}$$

利用凸性，仿照（6）中的论证，我们可以得到上界

$$\sup_k |K_k * f| \leq \mathbf{C} \cdot \sup_k |K_k' * f|,$$

因此，这样替换并无妨碍。

因此，我们将关注如何理解

$$K_k' * \varphi_{Q,N}. \tag{17}$$

我们快速指出一些缩放考量：由于

$$|n^d - (n-1)^d| \geq 2^{k(d-1)}$$

对于 $2^{k-1} < n \leq 2^{k+2}$，（17）在 $N \leq 2^{k(d-1)}$时变得平凡，因为此时和集中的每个元素

$$\{n^d : 2^{k-1} < n \leq 2^{k+2}\} + \{Qj : j \leq N/Q\}$$

有 O 个形如 $n^d + Qj$ 的表示。另一方面，由于 K_k' 支撑在 $[2^{dk+2}]$上，我们可以假设 $N \leq 2^{dk+2}$，因为与 K_k' 的卷积在由 $> 2^{dk+2}$分隔的区间上彼此独立。特别地，利用平移不变性，我们可以且将限制在 $|x| \leq \mathbf{C} \cdot 2^{dk}$，并假设

$$N \approx 2^{k(d-1+\delta)} \tag{18}$$

对于某个 $0 < \delta \leq 1$。

如果使用傅里叶反演，我们可以表示

$$(17) = \int \widehat{K_k'}(\beta) \cdot \widehat{\varphi_{Q,N}}(\beta) \cdot e(\beta x)\, d\beta. \tag{19}$$

为了确定 $\varphi_{Q,N}$ 的傅里叶变换，我们将 $Q\mathbb{Z}$ 的示性函数表示为指数和，

$$\mathbf{1}_{Q\mathbb{Z}}(n)=\frac{1}{Q}\sum_{A=1}^{Q}e(A/Q\cdot n),$$

并计算

$$\sum_{n}\frac{1}{Q}\sum_{A=1}^{Q}e(A/Q\cdot n)\cdot\varphi(n/N)\cdot e(-n\beta)=\frac{1}{Q}\sum_{A=1}^{Q}N\widehat{\varphi}(N(\beta-A/Q)) \tag{20}$$

对施瓦茨函数应用泊松求和

$$t\mapsto\frac{1}{Q}\sum_{A=1}^{Q}e(A/Q\cdot t)\cdot\varphi(t/N)\cdot e(-t\beta)$$

特别是，考虑到施瓦茨尾项，我们只关心

$$\beta\in\mathbb{Z}/Q\mathbb{Z}+O(N^{\mathbf{c}_{d,\delta}-1}),$$

与相反的情况一样

$$\Big|\sum_{n}\frac{1}{Q}\sum_{A=1}^{Q}e(A/Q\cdot n)\cdot\varphi(n/N)\cdot e(-n\beta)\Big|\leq\mathbf{C}_{d,\delta}\cdot N^{-100}$$

利用φ的施瓦茨衰减，参见（20）。因此，对于这样的 β，分解

$$\beta=A/Q+\eta,\quad|\eta|\leq\mathbf{C}\cdot N^{\mathbf{c}_{d,\delta}-1},$$

以及 $n=pQ+r$，我们发现

$$\begin{aligned}\beta n^{d}&=(A/Q+\eta)\cdot(pQ+r)^{d}\\&\equiv A/Q\cdot r^{d}+\eta\cdot(pQ)^{d}+O(|\eta|\cdot 2^{(d-1)k}\cdot Q)\mod 1,\end{aligned}$$

因此，对于这样的 β

$$\begin{aligned}\widehat{K_k'}(\beta)&=\sum_{n}\phi_k(n)\cdot e(-\beta n^{d})\\&=\sum_{pQ+r}\phi_k(pQ+r)\cdot e(-A/Q\cdot r^{d})\cdot e(-\eta\cdot(pQ)^{d})+O\Big(\frac{2^{k(d-1)}\cdot Q}{N^{1-\mathbf{c}_{d,\delta}}}\Big)\\&=\frac{1}{Q}\sum_{r=1}^{Q}e(-A/Q\cdot r^{d})\cdot\sum_{pQ}Q\cdot\phi_k(pQ)\cdot e(-\eta\cdot(pQ)^{d})+O\Big(\frac{2^{k(d-1)}\cdot Q}{N^{1-\mathbf{c}_{d,\delta}}}\Big),\end{aligned} \tag{21}$$

利用 ϕ 的光滑性。为了消去这些误差项，我们规定 $Q\leq 2^{k\delta/2}$，见（18），因此对于 $|\beta-A/Q|\leq\mathbf{C}\cdot N^{\mathbf{c}_{d,\delta}-1}$ 我们可以表示为

$$K_k'(\beta)=S(A/Q)\cdot\sum_{pQ}Q\cdot\phi_k(pQ)\cdot e(-(\beta-A/Q)\cdot(pQ)^{d})+\widehat{\mathcal{E}_k}(\beta),$$

其中 $S(A/Q)$ 是完全魏尔和，E_k 是傅里叶系数较小的误差项。具体来说：

$$S(A/Q):=\frac{1}{Q}\sum_{n\leq Q}e(-A/Q\cdot n^{d})=\frac{1}{Q}\sum_{m\leq Q}e(-A/Q\cdot m)\cdot|\{n\leq Q:n^{d}\equiv m\mod Q\}|$$

通过上界对 n^d 模 Q 的等分布性质进行了精确刻画，

$$|S(A/Q)| \le \mathbf{C}_\epsilon \cdot Q^{\epsilon-\frac{1}{d}}, \quad (A,Q)=1, \quad \epsilon>0; \tag{22}$$

见 [18]。并且，E_k 是一个可忽略的误差项，因为

$$\|\widehat{\mathcal{E}_k}\|_{L^\infty(\mathbb{T})} \le \mathbf{C}\cdot 2^{-k\delta/4}$$

（只要适当选取$\mathbf{c}_{d,\,\delta}$），便有

$$\|\mathrm{E}_k * \varphi_{Q,N}\|_{\ell^2} = \|\widehat{\mathcal{E}_k}\cdot\widehat{\varphi_{Q,N}}\|_{L^2(\mathbb{T})} \le \mathbf{C}\cdot 2^{-k\delta/4}\cdot\|\widehat{\varphi_{Q,N}}\|_{L^2(\mathbb{T})} = \mathbf{C}\cdot 2^{-k\delta/4}\cdot\|\varphi_{Q,N}\|_{\ell^2(\mathbb{Z}\,)};$$

在下文中，我们将不再考虑 E_k。

通过黎曼求和论证，比较

$$\begin{aligned} Q\cdot\phi_k(Qp)\cdot e(-(pQ)^d(\beta-A/Q)) = \int_p^{p+1} Q\cdot\phi(Qt)\cdot e(-(\beta-A/Q)\cdot(tQ)^d)\,dt \\ + O\Big(2^{-\mathbf{c}_{d,\delta}k}\cdot 2^{-k}\cdot Q\cdot(1+2^{-k}\cdot|Qp|)^{-100}\Big) \end{aligned}$$

我们将其近似到逐点误差为 $2^{-\mathbf{c}_{d,\,\delta}k}$的量级

$$\begin{aligned} \widehat{K_k'}(\beta) &= S(A/Q)\cdot\int\phi(t)\cdot e(-2^{dk}(\beta-A/Q)\cdot t^d)\,dt + O(2^{-\mathbf{c}_{d,\delta}k}) \\ &= S(A/Q)\cdot\int\phi'(s)\cdot e(-2^{dk}(\beta-A/Q)\cdot s)\,ds + O(2^{-\mathbf{c}_{d,\delta}k}), \qquad \phi'(s) := \frac{\phi(s^{1/d})}{ds^{1-1/d}} \\ &= S(A/Q)\cdot\widehat{\phi'}(2^{dk}(\beta-A/Q)) + O(2^{-\mathbf{c}_{d,\delta}k}) \end{aligned}$$

其中 ϕ' 也是施瓦茨函数，见（16）。因此

$$\begin{aligned} \text{“}=\text{”}\quad & \frac{1}{Q}\sum_{A\le Q} S(A/Q)\int N\widehat{\varphi}(N(\beta-A/Q))\cdot\widehat{\phi'}(2^{dk}(\beta-A/Q))\cdot e(\beta x)\,d\beta, \\ &= \frac{1}{Q}\sum_{A\le Q} e(A/Qx)\cdot S(A/Q)\cdot\Phi(x), \end{aligned} \tag{19}$$

这里我们将其合并

$$\Phi(x) := \int\varphi((x-2^{dk}s)/N)\cdot\phi'(s)\,ds$$

以便

$$\hat{\Phi}(\beta) = N\hat{\varphi}(N\beta)\cdot\widehat{\phi'}(2^{dk}\beta),$$

因此 $\|\Phi\|_{\ell^2(\mathbb{Z})} \approx \frac{N}{2^{dk/2}}$。求和可得

$$\begin{aligned} \|K_k'*\varphi_{Q,N}\|_{\ell^2(\mathbb{Z})}^2 \text{“}=\text{”} &\sum_x\Big|\frac{1}{Q}\sum_{A\le Q}e(A/Qx)\cdot S(A/Q)\Big|^2\cdot|\Phi(x)|^2 \\ &= \frac{1}{Q^2}\sum_{A,B\le Q} S(A/Q)\cdot\overline{S(B/Q)}\cdot\sum_x e((A/Q-B/Q)x)\cdot|\Phi(x)|^2 \\ &= \frac{1}{Q^2}\sum_{A,B\le Q} S(A/Q)\cdot\overline{S(B/Q)}\cdot\widehat{|\Phi|^2}(A/Q-B/Q). \end{aligned} \tag{23}$$

由于

$$\widehat{|\Phi|^2} = \hat{\Phi} * \hat{\Phi}^*, \quad \text{where} \quad g^*(x) := \overline{g(-x)}$$

基本支撑在 $\{|\xi| \le \mathbf{C} \cdot N^{-1}\}$ 内，因此我们有

$$\widehat{|\Phi|^2}(A/Q - B/Q) = \delta_{A=B} \cdot \|\Phi\|^2_{\ell^2(\mathbb{Z})} + O((N/Q)^{-100}) \tag{24}$$

由于 $A \neq B$, $|A/Q - B/Q| \ge Q^{-1} \gg N^{-1}$。将（24）代入（23），得到

$$\begin{aligned} \|K_k' * \varphi_{Q,N}\|^2_{\ell^2(\mathbb{Z})} \text{“} &= \text{”} \frac{1}{Q^2} \sum_{A,B \le Q} S(A/Q) \cdot \overline{S(B/Q)} \cdot \delta_{A=B} \cdot \|\Phi\|^2_{\ell^2(\mathbb{Z})} \\ &= \frac{1}{Q^2} \sum_{A \le Q} |S(A/Q)|^2 \cdot \|\Phi\|^2_{\ell^2(\mathbb{Z})}. \end{aligned}$$

根据华的估计（22），利用 Q 是素数这一事实，我们得到

$$\frac{1}{Q} \sum_{A \le Q} |S(A/Q)|^2 = \frac{1}{Q} + \frac{1}{Q} \sum_{A \le Q-1} |S(A/Q)|^2 \le \mathbf{C}_\epsilon \cdot (1/Q + Q^{\epsilon - 2/d})$$

由此可得

$$\begin{aligned} \|K_k' * \varphi_{Q,N}\|_{\ell^2(\mathbb{Z})} &\le \mathbf{C}_\epsilon \cdot Q^{\epsilon - 1/d} \cdot Q^{-1/2} \cdot \frac{N}{2^{dk/2}} \\ &= \mathbf{C}_\epsilon \cdot Q^{\epsilon - 1/d} \cdot (N/2^{dk})^{1/2} \cdot (N/Q)^{1/2} \\ &\le \mathbf{C}_\epsilon \cdot Q^{\epsilon - 1/d} \cdot (N/2^{dk})^{1/2} \cdot \|\varphi_{Q,N}\|_{\ell^2(\mathbb{Z})} \end{aligned}$$

前置因子 $(N/2^{dk})^{1/2}$ 来自缩放考量；如果我们关心的是一个与尺度无关的估计，则可得到以下界

$$\|K_k' * \varphi_{Q,N}\|_{\ell^2(\mathbb{Z})} \le \mathbf{C}_\epsilon \cdot Q^{\epsilon - 1/d} \cdot \|\varphi_{Q,N}\|_{\ell^2(\mathbb{Z})}.$$

特别是，在定量意义上，下界

$$\|K_k' * \varphi_{Q,N}\|_{\ell^2(\mathbb{Z})} \ge \delta \cdot \|\varphi_{Q,N}\|_{\ell^2(\mathbb{Z})}$$

自动　地迫使$\varphi_{Q,\ N}$通过估计得到“算术复杂度”的一个界

$$Q \le \mathbf{C}_\epsilon \cdot \delta^{-d-\epsilon}.$$

特别是，我们得出以下启发式结论：

启发式 2.1。——唯一的障碍是

$$\|K_k' * f\|_{\ell^2(\mathbb{Z})} \ll \|f\|_{\ell^2(\mathbb{Z})}$$

属于“低算术复杂度”的考量。

2.2. 要点

通过应用 Weyl 引理（其中特殊情形如下所述），Bourgain 得以将此前的启发式论证 2.1 严格化，并得出上述例子范围具有典型性的结论：若我们令

$$\Pi_k(\beta) := \sum_{(A,Q)=1, Q\le 2^{\mathbf{c}k}} \widehat{\chi}(2^{(d-\mathbf{c})k}(\beta - A/Q))$$

对于一个满足以下条件的施瓦茨函数 χ

$$\mathbf{1}_{[-1/4,1/4]} \le \widehat{\chi} \le \mathbf{1}_{[-1/2,1/2]}$$

then

$$\widehat{K_k}(\beta) = \widehat{K_k}(\beta)\cdot \Pi_k(\beta) + O(2^{-\mathbf{c}'k}), \tag{25}$$

对于 K'_k，情况也类似。特别是，只要 $\Pi_k(\beta) \neq 1$，Weyl 引理的结论就必然成立。

引理 2.2 （Weyl 引理，特殊情形）。— 假设 $|\beta - a/q| \le {}_q{}^{1}{}_{\cdot N^{d-\mathbf{c}}}$ ~~wi~~ th

$$N^{\mathbf{c}} \le q \le N^{d-\mathbf{c}}.$$

那么存在某个 $\mathbf{c}_d > 0$，使得

$$|\frac{1}{N}\sum_{n\le N} e(-\beta n^d)| \le \mathbf{C}_d \cdot N^{-\mathbf{c}_d}.$$

此时，沿用前一个例子中的推理，便可得到物理空间中的近似

$$K'_k \quad \text{“}=\text{”} \quad L'_k := \sum_{(A,Q)=1,\ Q\le 2^{\mathbf{c}k}} S(A/Q)\cdot \mathrm{Mod}_{A/Q}(\chi_{(d-\mathbf{c})k} * \phi'_{dk})$$

在此条件下

$$\|\widehat{K'_k - L'_k}\|_{L^\infty(\mathbb{T})} \le \mathbf{C}_d \cdot 2^{-\mathbf{c}_d k} \tag{26}$$

因此，极大函数在 $\ell^2(\mathbb{Z})$ 上有界

$$\|\sup_k |(K'_k - L'_k) * f\|^2_{\ell^2(\mathbb{Z})} \le \sup_\beta \sum_k |\widehat{K'_k}(\beta) - \widehat{L'_k}(\beta)|^2 \cdot \|f\|^2_{\ell^2(\mathbb{Z})} \le \mathbf{C}_d \cdot \|f\|^2_{\ell^2(\mathbb{Z})}, \tag{27}$$

按照 (12) 中的论证方式，并在最后一步不等式中代入定量界 (26)。

依照启发式 2.1，将 L'_k 按高斯和的近似水平集分解，并寻求 s 在 $\ell^2(\mathbb{Z})$-范数下的极大函数

$$\sup_{k\ge \mathbf{C}s} |L'_{k,s} * f|,$$

其中

$$L'_{k,s} := \sum_{A/Q\in \mathcal{R}_s} S(A/Q)\cdot \mathrm{Mod}_{A/Q}(\chi_{(d-\mathbf{c})k} * \phi'_{dk})$$

for

$$\mathcal{R}_s := \{(A,Q)=1,\ 2^{s-1}\le Q<2^s\}: \tag{28}$$

一个界

$$\sup_k |L'_k * f| = \sup_k |\sum_{s\le \mathbf{c}k} L'_{k,s} * f| \le \sum_{s=1}^{\infty} \sup_{k\ge \mathbf{C}s} |L'_{k,s} * f|.$$

稍作巧妙处理，利用普朗歇雷定理从中抽取一个几何衰减的前置因子，

$$L'_{k,s}(x) \quad \text{“}=\text{”} \quad 2^{-\mathbf{c}_d s}\cdot \sum_{A/Q\in\mathcal{R}_s} \mathrm{Mod}_{A/Q}(\chi_{(d-\mathbf{c})k} * \phi'_{dk})(x)$$

只需证明以下极大不等式（可能需要稍微改变 χ的选取）：

$$\| \sup_{k\ge \mathbf{C}s} |\Big(\sum_{A/Q\in\mathcal{R}_s} \mathrm{Mod}_{A/Q}\chi_k\Big) * f|\|_{\ell^2(\mathbb{Z})} \le \mathbf{C}_\epsilon \cdot 2^{\epsilon s}\cdot \|f\|_{\ell^2(\mathbb{Z})},\quad \epsilon>0;$$

通过对平移取平均，并利用$\{\chi_k:\ k\ge \mathbf{C}s\}$在物理尺度 $2^{\mathbf{C}s}$上的光滑性，只需证明相应的实变量不等式:

$$\| \sup_{k\ge \mathbf{C}s} |\Big(\sum_{A/Q\in\mathcal{R}_s} \mathrm{Mod}_{A/Q}\chi_k\Big) * f|\|_{L^2(\mathbb{R})} \le \mathbf{C}_\epsilon \cdot 2^{\epsilon s}\cdot \|f\|_{L^2(\mathbb{R})},\quad \epsilon>0;$$

最后，利用 $\mathbb{R}$的伸缩不变性，问题归结为证明以下多频极大估计，见 [9]:

命题 2.3。——假设 $\Theta:=\{\theta_1,\dots,\theta_N\}$ 两两1分离，

$$i.e.\quad |\theta_i-\theta_j|>1,\quad i\ne j.$$

Then

$$\|\mathcal{M}_\Theta f\|_{L^2(\mathbb{R})} := \| \sup_{k\ge \mathbf{C}} |\sum_{n\le N} (\mathrm{Mod}_{\theta_n}\chi_k) * f|\|_{L^2(\mathbb{R})} \le \mathbf{C}_\epsilon \cdot N^{\epsilon}\cdot \|f\|_{L^2(\mathbb{R})},\quad \epsilon>0. \tag{29}$$

命题 2.3 的证明（我们马上就会给出，右侧（29）的界为 $\log^2 N$）结合了调和分析、概率论和巴拿赫空间几何中的思想，是一项创新成果，并进一步应用于逐点遍历理论中的问题 [10]; [13]; [14] 以及时频分析中的问题，例如 [15]; [24]。另一方面，从某些角度看，这种证明技巧受到很大限制：控制 L^2上的极大函数的方法毕竟有限，下面我们将对此进行探讨。

3. 多频问题

3.1. 初步观察

接下来，我们引入以下记号

$$\Xi_k f := \sum_{n\le N} (\mathrm{Mod}_{\theta_n}\chi_k) * f, \tag{30}$$

从而我们可以表示为

$$\mathcal{M}_\Theta f = \sup_k |\Xi_k f|;$$

这里与上文一样， χ是一个满足以下条件的施瓦茨函数

$$\mathbf{1}_{[-1/4,1/4]} \le \hat{\chi} \le \mathbf{1}_{[-1/2,1/2]}$$

虽然 Ξ_k 具有振荡核，但它们具有自然的投影结构，因为

$$\Xi_k\Xi_l = \Xi_l, \quad k \ge l+2,$$

如转到傅里叶空间即可看出，见下文（31）；为避免不必要的技术细节，以下我们将尺度集稀疏化为奇偶类，并只关注其中一个类，因此每当 $k > k'$时，必有 $k \ge k' + 2$。

由于证明命题2.3 是一个基于 L^2的问题，为了更好地理解这些卷积算子，我们转到傅里叶空间并计算

$$\widehat{\Xi_k f}(\xi) = \sum_{n\le N} \widehat{\chi_k}(\xi-\theta_n)\cdot \hat{f}(\xi) \tag{31}$$

于是

$$\widehat{\Xi_k}(\xi) = \sum_{n\le N} \widehat{\chi_k}(\xi-\theta_n),$$

将算子与其核视为一体后，我们也可以表示为

$$\begin{aligned}
\Xi_k f(x) &= \sum_{n\le N} e(\theta_n x)\int \widehat{\chi_k}(\xi)\hat{f}(\xi+\theta_n)e(\xi x)\,d\xi \\
&= \sum_{n\le N} e(\theta_n x)\cdot\Big(\chi_k * (\mathrm{Mod}_{-\theta_n} f)\Big)(x) \\
&= \sum_{n\le N} e(\theta_n x)\cdot\Big(\chi_k * (\chi * \mathrm{Mod}_{-\theta_n} f)\Big)(x) \\
&=: \sum_{n\le N} e(\theta_n x)\cdot(\chi_k * f_{\theta_n})(x), \qquad (32)
\end{aligned}$$

利用 $k \ge \mathbf{C}$ 这一事实，并通过一个关于傅里叶变换的简短论证，得到再生恒等式

$$\chi_k * \chi = \chi.$$

转而采用涉及 $\{f_{\theta_n}\}$ 的表述，其优点在于，卷积与 χ_k 的平滑效应已从振荡指数函数 $\{e(\theta_n x) : n\}$ 中“因式分解”出来。特别是，启发式地说，在有界长度的区间 $\mathbf{C}$ 上，当 k 变大时

只有指数函数应当发生变化：如果 $|I| = \mathbf{C}$ 是一个长度适当的区间，那么每当 $x \in I$ 和 $2^k \gg \mathbf{C}$

$$I \ni x \mapsto \sum_{n \le N} e(\theta_n x) \cdot \phi_k * f_{\theta_n}(x) \quad \text{“}=\text{”} \quad \sum_{n \le N} e(\theta_n x) \cdot \phi_k * f_{\theta_n}(x_I) \tag{33}$$

对于任意 x_I。特别地，如果我们将 $\mathbb{R}$ 细分为（双进）区间 $\{I : |I| = \mathbf{C}\}$，那么在每个区间上都可以估计

$$\int_I |\Xi_k f(x)|^2 \, dx \quad \text{“}=\text{”} \quad \min_{x_I \in I} \int_I |\sum_n e(\theta_n x) \cdot \chi_k * f_{\theta_n}(x_I)|^2 \, dx \\ \le \mathbf{C} \cdot \min_{x_I \in I} \sum_{n \le N} |\chi_k * f_{\theta_n}(x_I)|^2 \cdot |I| \tag{34}$$

这可通过对下式进行界定看出

$$\begin{aligned} \| \sum_{n \le N} e(\theta_n x) \cdot a_n \|_{L^2(I)}^2 \le \| \sum_{n \le N} e(\theta_n x) \cdot a_n \cdot \psi_I(x) \|_{L^2(\mathbb{R})}^2 &= \| \sum_{n \le N} a_n \cdot \widehat{\psi_I}(\xi - \theta_n) \|_{L^2(\mathbb{R})}^2 \\ = \sum_{n \le N} a_n \overline{a_m} \cdot \langle \widehat{\psi_I}(\cdot - \theta_n), \widehat{\psi_I}(\cdot - \theta_m) \rangle &= \sum_{n \le N} |a_n|^2 \cdot \|\widehat{\psi_I}\|_{L^2(\mathbb{R})}^2 \\ \le \mathbf{C} \cdot \sum_{n \le N} |a_n|^2 \cdot |I|, & \end{aligned}$$

对于 $\mathbf{1}_I \le |\psi_I| \le \mathbf{C} \cdot (1 + \operatorname{dist}(\cdot\ I)/|I|)^{-100}$，其傅里叶变换紧支撑于 $[-1/2, 1/2]$ 内；这一支撑约束保证

$$\psi_I(\xi - \theta_n) \cdot \psi_I(\xi - \theta_m) \equiv 0, \quad n \ne m,$$

且 s ince $|I| \ge \mathbf{C}$，满足不确定性原理，因此可以选取这样的 ψ_I .

为了寻求一致性，若我们令

$$\mathcal{F}_\Theta(x)^2 := \mathcal{F}_{\Theta, \mathbf{C}}(x)^2 := \sum_{n \le N} \sup_{k \ge \mathbf{C}} |\chi_k * f_{\theta_n}(x)|^2, \tag{35}$$

那么我们就有一个一致的范数界

$$\sup_{k \ge \mathbf{C}} \|\Xi_k f\|_{L^2(I)} \le \mathbf{C} \cdot \min_{x_I \in I} \mathcal{F}_\Theta(x_I) \cdot |I|^{1/2} \le \|\mathcal{F}_\Theta\|_{L^2(I)}, \tag{36}$$

而我们的任务是控制

$$\|\mathcal{M}_\Theta f\|_{L^2(\mathbb{R})} = \Big(\sum_{|I| = \mathbf{C}} \| \sup_{k \ge \mathbf{C}} |\Xi_k f| \|_{L^2(I)}^2 \Big)^{1/2}, \tag{37}$$

前面的计算促使我们将实数轴划分为长度“较小”的区间，并分别处理 $\mathbf{M}_\Theta f$ 在每个区间上的贡献。因此，问题归结为控制 L^2 上的一个上确界：我们可以局部化，并

通过以下界来处理每个单独的 Ξ_k 的贡献

$$\|\mathcal{F}_\Theta\|_{L^2(\mathbb{R})}^2 = \sum_{n\le N}\int \sup_{k\ge \mathbf{C}} |\chi_k * f_{\theta_n}(x)|^2 \le \mathbf{C}\cdot\sum_{n\le N}\int |f_{\theta_n}(x)|^2\,dx = \mathbf{C}\cdot\sum_{n\le N}\int |\widehat{f_{\theta_n}}(\xi)|^2\,d\xi$$

$$= \mathbf{C}\cdot\sum_{n\le N}\int |\hat\chi(\xi)|^2\cdot|\hat f(\xi+\theta_n)|^2\,d\xi = \mathbf{C}\cdot\int\sum_{n\le N}|\hat\chi(\xi-\theta_n)|^2\cdot|\hat f(\xi)|^2\,d\xi$$

$$\le \mathbf{C}\cdot\sup_\xi\sum_{n\le N}|\hat\chi(\xi-\theta_n)|^2\cdot\|\hat f\|_{L^2(\mathbb{R})}^2 \le \mathbf{C}\cdot\|f\|_{L^2(\mathbb{R})}^2, \tag{38}$$

利用频率 $|\theta_n-\theta_m|>1,\ n\neq m$ 的分离性，以及第一个不等式中的 Hardy–Littlewood极大函数。我们的任务是通过 $\mathbf{M}_\Theta$，从对 $\{\Xi_k : k\}$ 的一致 控制过渡到同时控制。这是一项常见的任务，但往往受到限制 —— 我们先停下来探讨一下。

3.2. 对 L^2 上确界的估计

设 $\{F_k\}\in L^2(X)$ 是测度空间上的一族函数，我们希望控制

$$\|F_*\|_{L^2(X)} := \|\sup_k |F_k|\|_{L^2(X)}. \tag{39}$$

据我所知，控制 F_* on L^2 基本上有四种方法：

– 鞅／停时方法，例如利用鞅理论证明 Doob 极大不等式时所用的方法，或与之密切相关的 Hardy–Littlewood极大不等式；– 半群方法，例如 Hopf–Dunford–Schwartz 极大定理所用的方法；该定理的一个特殊情形可推出极大函数 $\sup_t |e^{t\triangle}f|$ 的与维数无关的界；– TT^* 正交性方法，其中上确界 F_* 由某个特定的线性算子实现，

$$T\{F_k\} := \sum_t \mathbf{1}_{E_k}F_k$$

对于 $\{E_k\}$ 的一个不相交划分，然后将 T 与其伴随算子复合，以高效计算

$$\|T\|_{L^2(X)\to L^2(X)} = \|TT^*\|_{L^2(X)\to L^2(X)}^{1/2};$$

这种技术常见于振荡积分情形；并且

– 熵论证，利用映射 $k\mapsto F_k(x)$ 中的残余光滑性来控制 F_*。

在这四种方法中，平均值 $\{\Xi_k : k\}$ 的振荡性排除了直接采用第一种方法进行论证的可能；第一种方法赋予零频率（期望）特殊地位；恒等式的严重失效

$$\Xi_k\Xi_l \neq \Xi_{k+l}$$

排除了第二种方法。

As 对于 TT^* 方法，如果我们将上确界线性化，并考虑算子 r

$$Tf(x) = \sum_k \mathbf{1}_{E_k}(x) \cdot \int \sum_{n \leq N} e(\theta_n x) \int \chi_k(x-y) e(-\theta_n y) f(y)\, dy,$$

那么，对偶算子 T^* 可表示为

$$T^* g(x) = \sum_r \sum_{n \leq N} e(\theta_n y) \cdot \int e(-\theta_n x) \cdot (g \cdot \mathbf{1}_{E_r})(x) \overline{\chi_k}(x-y)\, dx,$$

而组合运算并没有真正带来任何收益。

因此，我们将注意力转向熵方法，以求对 L^2 上的上确界进行界定。

4. 源自 Bourgain 工具箱：熵方法

本节回顾的是 Bourgain 在 [9] 时已经完全掌握的内容；参见 [4]； [5]； [7]； [8]，甚至还可参见 §3 of [9] 中的典型例子，以及 §6 of[35]中的出色总结。特别是，我想，Bourgain 从上述第 3.1 小节中获得的信息，足以直接引导他得出下文第 5节的内容。初读时，这种方法在研究 $\mathcal{M}_\Theta$ 时似乎颇具魔力 [2], ；至少对我而言，阅读 [36], 后，我希望读者在充分消化以下内容后，能够理解 Bourgain 得出其论证时的思路。

熵方法背后的基本机制，是利用参数空间中的“大小”和“光滑性”，或者更确切地说，“黏性”，来控制上确界。就我们当前的问题而言，我们通过（ Ξ_k 36）对每个平均值进行一致控制，并寻找某种光滑性／黏性的概念来补足这种一致性。

为了展示这种相互作用，我们回顾以下例子。

引理 4.1 （索伯列夫嵌入引理）。— 假设 I 是一个区间，并且对于几乎每个 x，F（x，·）绝对连续，且具有一个 L^2 密度。那么有如下逐点估计：

$$F_I(x) := \sup_{t \in I} |F(x,t)| \leq \mathbf{C} \cdot |F(x,t_I)| + \mathbf{C} \cdot \left(\int_I |F(x,t)|^2\, dt \right)^{1/4} \cdot \left(\int_I |\partial_t F(x,t)|^2\, dt \right)^{1/4}$$

对于任意 $t_I \in I$。特别是，如果

$$\sup_{t \in I} \|F(x,t)\|_{L^2_x} \leq A \quad \textit{and} \quad \sup_{t \in I} \|\partial_t F(x,t)\|_{L^2_x} \leq a \tag{40}$$

then

$$\| \sup_{t \in I} |F(x,t)| \|_{L^2_x} \leq \mathbf{C} \cdot \left(A + (Aa|I|)^{1/2} \right)$$

证明。——对于任意 $t \in I$，我们可以估计

$$F(x,t)^2 = F(x,t_I)^2 + \int_{[t_I,t]} \partial_s \Big(F(x,s)^2 \Big)\, ds,$$

so

$$\begin{aligned} |F(x,t)|^2 &\le |F(x,t_I)|^2 + 2\int_I |F(x,t)| \cdot |\partial_t F(x,t)|\, dt \\ &\le |F(x,t_I)|^2 + 2\left(\int_I |F(x,t)|^2\, dt\right)^{1/2} \cdot \left(\int_I |\partial_t F(x,t)|^2\, dt\right)^{1/2} \end{aligned} \tag{41}$$

(41) 的右侧与 t无关，因此我们可以对 t 取上确界，作用于 (41) 的左侧，然后对 x积分，并应用柯西–施瓦茨处理基于 L^2的 t平均值。

□

虽然引理 4.1的证明非常简短，但它出乎意料地稳健，在研究如下形式的极大乘子算子时非常有用

$$\sup_t |(\hat{f} \cdot \mathbf{m}(t\cdot))^\vee|$$

对于有界的 $m \in C^1(\mathbb{R} \setminus \{0\})$，参见 [4] 的引理 3。

将这一论证离散化会很有帮助：对每个 $v \ge 1$，定义

$$\Lambda_v := \left(2^{-v} \cdot \mathbb{Z}\right) \cap I$$

并将 $t \in \Lambda_v$, $\rho(t) \in \Lambda_{v-1}$ 的父节点定义为满足以下条件的最小元素

$$B(t, 2^{-v}) \cap B(\rho(t), 2^{1-v}) \neq \emptyset, \qquad B(x,s) := \{y : |x-y| < s\}.$$

由于在 $t \mapsto F(x,t)$ 中 x-几乎处处连续，要研究 F_I，只需对其进行界定

$$\sup_{t \,\in\, \bigcup_{v\ge 1} \Lambda_v} |F(x,t)|;$$

由单调收敛可知，只需对有限子集 $T \subset \bigcup_{v\ge 1} \Lambda_v$ 一致地估计

$$F_T(x) := \sup_{t\in T} |F(x,t)|.$$

为此，对于每个 $t \in T$，我们可以作望远镜求和

$$t = (t - \rho(t)) + (\rho(t) - \rho^2(t)) + \cdots + \rho_0(t)$$

其中 ρ^j 是 ρ 的第 j 次复合，而 $\rho_0(t)$ 是适当的复合，使得 $\rho_0(t) \in \Lambda_{v_0}$，其中 v_0 将在下文确定。

注意，由于 T 是有限的，到达代表元 $\rho_0 \in \Lambda_{v_0}$ 所需的增量次数有统一上界。我们对其进行估计

$$\begin{aligned} F_T(x) &\le \sup_{t\in\Lambda_{v_0}} |F(x,t)| + \sum_{v>v_0} \sup_{t\in\Lambda_v} |F(x,t) - F(x,\rho(t))| \\ &\le \Big(\sum_{t\in\Lambda_{v_0}} |F(x,t)|^2 \Big)^{1/2} + \sum_{v>v_0} \Big(\sum_{t\in\Lambda_v} |F(x,t) - F(x,\varrho(t))|^2 \Big)^{1/2}, \end{aligned}$$

注意到所有求和实际上都是有限的，先取 L^2_x-范数，再对 $v_0 \geq 0$ 进行优化，从而得到所需的上界：

$$A \cdot |\Lambda_{v_0}|^{1/2} + a \cdot \sum_{v > v_0} |\Lambda_v|^{1/2} \cdot 2^{-v} \leq \mathbf{C} \cdot \Big(A + (Aa|I|)^{1/2}\Big);$$

见（40）。

在这两个论证中，我们都依赖于映射 $t \mapsto F(x,t)$ 的光滑性。不过，实际上，我们依赖的是以下各项贡献的衰减：

$$\Lambda_v \ni t \mapsto |F(x,t) - F(x,\rho(t))| \tag{42}$$

随着 v 增大，以及受控的熵估计

$$|\Lambda_v| \leq \mathbf{C} \cdot 2^v \cdot |I|;$$

从度量视角来看，这个估计衡量的是I 中的元素在多大程度上彼此贴合——“粘在一起”——其尺度为 2^{-v}。诸如此类的估计

$$\sup_{t \in I} \|\partial_t F(x,t)\|_{L^2_x} \leq a$$

使我们能够以 L^2平均意义刻画（42）中的小量。不过，我们也可以用类似的望远镜式机制，更直接地逐点逼近 $\{F(x,\ t)\colon\ t \in I\}$。

对于如上所述的 T，考虑集合

$$X(x) := X_T(x) := \Big\{F(x,t) : t \in T\Big\}, \tag{43}$$

并且对于每个满足 $v \cdot \mathrm{diam}(X(x))$ 的 v，定义 $\Lambda_v(x) \subset T$ 为一个时间集合 t，使得

$$X(x) \subset \bigcup_{t \in \Lambda_v(x)} B\Big(F(x,t), 2^{-v}\Big) \tag{44}$$

满足约束条件 $|\Lambda_v(x)|$ 为最小；其基数本质上就是该集合的 2^{-v}-熵 。

现在，令 V 足够大，使得 T 中的每个元素都以 $> 2^{1-V}$ 分隔，从而$T = \Lambda_V(x)$。并定义 $t \in \Lambda_v(x)$ 的父节点 $\varrho(t) \in \Lambda_{v-1}(x)$ 为满足条件的最小元素，使得

$$B\Big(F(x,t), 2^{-v}\Big) \cap B\Big(F(x,\varrho(t)), 2^{1-v}\Big) \neq \emptyset. \tag{45}$$

对于任意 $s \in T$，我们同样可以估计

$$\begin{aligned} F_T(x) &\leq |F(x,s)| + \sum_v \sup_{t \in \Lambda_v(x)} |F(x,t) - F(x,\varrho(t))| \\ &\leq |F(x,s)| + \sum_v \Big(\sum_{t \in \Lambda_v(x)} |F(x,t) - F(x,\varrho(t))|^2 \Big)^{1/2} \\ &\leq |F(x,s)| + \mathbf{C} \cdot \sum_v 2^{-v} \cdot |\Lambda_v(x)|^{1/2}, \end{aligned} \tag{46}$$

正如我们可以估计

$$|F(x,t)-F(x,\rho'(t))|<2^{-v}+2^{1-v}<2^{2-v}$$

对每个 $t\in\Lambda_v(x)$ 通过 (45。改变视角并给出界限会更方便

$$|\Lambda_v(x)|\le N_{2^{-v}}(x) \tag{47}$$

其中 $N_\lambda(x):=\sup\Big\{K:$ 存在一个时间序列

$$t_0<t_1<\cdots<t_K:|F(x,t_i)-F(x,t_{i-1})|>\lambda\Big\}$$

是一个所谓的（贪心）跳跃计数函数，其高度为 $\lambda>0$，用于衡量在尺度 λ 下，$\{F(x,t):t\}$“聚集在一起”的程度：

$$N_\lambda(x)<\infty \text{ for all } \lambda>0 \iff \{F(x,t):t\} \text{ converges}$$
$$\iff \{F(x,t):t\} \text{ “stick together” at all scales.}$$

为证明 (47)，对左侧取上界、对右侧取下界，使用集合的2^{1-v}-熵：集合中最大的、彼此相距至少 2^{1-v}的点集的大小，点位于$\{F(x,t):\,t\in I\}$之内。

反向界

$$N_{2^{-v}}(x)\le|\Lambda_{v+1}(x)|$$

更简单，因此这一变化在定量上并无损失，正如事实上

$$\sum_v 2^{-v}\cdot|\Lambda_v(x)|^{1/2}\le\sum_v 2^{-v}\cdot N_{2^{-v}}(x)^{1/2}\le\mathbf{C}\cdot\sum_v 2^{-v}\cdot|\Lambda_v(x)|^{1/2}.$$

在许多特殊例子中，我们能够证明一个一致界

$$\sup_v\|2^{-v}\cdot N_{2^{-v}}^{1/2}\|_{L^2}\le\mathbf{C}\cdot A, \tag{48}$$

这表明，在 L^2 平均意义下

$$N_{2^{-v}}\quad\text{“}\le\text{”}\quad\mathbf{C}\cdot A^2\cdot 2^{2v}$$

也就是说，要用给定半径的球覆盖数据集 $\{F(x,\ t):\ t\}$，其代价按二次方增长。以下例子具有代表性。

*E*熵论例子一。——考虑（离散时间的）平均算子，

$$F(x,t)=\mathbb{E}_k f(x)\cdot\mathbf{1}_{[2^k,2^{k+1})}(t) \tag{49}$$

其中

$$\mathbb{E}_k f(x):=\sum_{|I|=2^k\text{ dyadic}}\Big(\frac{1}{|I|}\int_I f(t)\,dt\Big)\cdot\mathbf{1}_I(x) \tag{50}$$

是条件期望算子，它将函数投影到由二进区间 σ生成的 $\{2^k\cdot\ [n,n+1):\ \ n\in\mathbb{Z}\}$-代数上。定义中嵌入的“停时”结构 $N_{2^{-v}}$ 使人能够巧妙地运用二进调和分析中的方法——其背后实际上是鞅方法——来证明（48）。

熵示例二。——在某种程度上

$$\mathbb{E}_k f \quad \text{“}=\text{”} \quad \chi_k * f,$$

由于这两个算子都会在空间尺度 2^k 上进行“模糊”处理，舍弃低于此阈值的“精细尺度”信息，同时保留可在高于此空间阈值处检测到的“粗略尺度”性质，因此可以将平方函数论证与进一步的正交性论证结合起来，特别是得到如下定量界

$$\|\mathbb{E}_k \psi_l - \chi_k * \psi_l\|_{L^2(\mathbb{R})} \leq \mathbf{C} \cdot 2^{-|k-l|/2} \cdot \|\psi_l\|_{L^2(\mathbb{R})}, \quad \psi_l := \chi_l - \chi_{l-1}, \tag{51}$$

将（48）推广到

$$F(x,t) = f * \chi_k(x) \cdot \mathbf{1}_{[2^k, 2^{k+1})}(t), \tag{52}$$

类似地，也可以将 χ 替换为任何其他满足 $\hat{\chi}(0) = 1$ 的施瓦茨函数。这些思想最早见于 [20]。

4.1. 熵的跳跃计数方法

尽管一致估计（48）先验上不足以控制对 $t \in T$的完整上确界，但这一熵论证相较于平凡估计带来了显著的加强

$$\|F_T\|_{L^2_x} \leq \|S_T\|_{L^2_x} \leq |T|^{1/2} \cdot A,$$

其中我们设

$$S_T(x)^2 := \sum_{t \in T} |F(x,t)|^2.$$

特别地，对于任意 $t \in T$，

$$\begin{aligned} F_T(x) &\leq |F(x,t)| + \mathbf{C} \cdot \sum_v 2^{-v} \cdot N_{2^{-v}}(x)^{1/2} \\ &\leq |F(x,t)| + \frac{S_T(x)}{|T|^{1/2}} + \sum_{v: \frac{S(x)}{|T|^{1/2}} \leq 2^{-v} \leq 2 \cdot S(x)} 2^{-v} \cdot N_{2^{-v}}(x)^{1/2} \end{aligned} \tag{53}$$

因此，实质上，一致界 (48) 蕴含[(2)]

$$\|F_T\|_{L^2_x} \quad \text{“}\leq\text{”} \quad \mathbf{C} \cdot \log |T| \cdot A. \tag{55}$$

事实上，正如下文将会看到的，(55) 通常成立，且带有对数^2 $|T|$ 前因子 .

---

(2)这个估计与抽象希尔伯特空间中的*Rademacher–Menshov* 不等式之间存在一种自然的比较；该不等式也指出

$$\| \sup_{n \leq N} |F(x,n)| \|_{L^2} \leq \mathbf{C} \cdot \log N \cdot A \tag{54}$$

在对函数 $\{F(\cdot\ n) : n\}$施加正交性约束时。这里的类比体现在证明层面，类似于勒贝格积分与黎曼积分之间的关系：熵界按照数据的像来组织数据 $\{F(x,\ n) : n\}$，而拉德马赫–门绍夫不等式则通过类似地按照时间参数 $n \in [N]$的定义域来组织数据，从而得到证明。

4.2. 变差简介

正如对（54）的启发式论证存在困难所表明的，见（53），一个主要问题是，一般而言，我们无法期待在

$$x \mapsto \sup_{v} \; 2^{-v} \cdot N_{2^{-v}}(x)^{1/2},$$

参见 [22]或 [33]。

为了解决这一问题，我们不再保留幂 $1/2 \to 1/r,\ r > 2$，而是引入所谓的 $\{F(x,t) : t \in I\}$ 的 r-变差

$$\mathcal{V}^r(x) := \mathcal{V}^r_F(x) := \sup \Big(\sum_i |F(x,t_i) - F(x,t_{i-1})|^r \Big)^{1/r}, \tag{56}$$

其中，上确界取遍 I 内所有有限递增子序列。与跳跃计数函数不同，r-变差算子满足三角不等式，这一点至关重要，

$$\mathcal{V}^r_{F+G} \le \mathcal{V}^r_F + \mathcal{V}^r_G,$$

并且可以对其作如下估计

$$\sup_{v} \; 2^{-v} \cdot N_{2^{-v}}(x)^{1/r} \le \mathcal{V}^r(x),$$

这很重要，因为 V^r 算子通常具有强大的 L^2-理论。特别是，若 $|T| = N$，使得对所有 λ都有 $N_\lambda \le N$，则我们可以估计

$$2^{-v} \cdot N_{2^{-v}}^{1/2} \le 2^{-v} \cdot N_{2^{-v}}^{1/r} \cdot N_{2^{-v}}^{1/2-1/r} \le N^{1/2-1/r} \cdot \mathcal{V}^r \tag{57}$$

若令 $r = 2 + \mathbf{c}\log N$，则可消去因子 $N^{1/2-1/r}$，最终得到估计

$$2^{-v} \cdot N_{2^{-v}}^{1/2} \le \mathbf{C} \cdot \mathcal{V}^r, \qquad r = 2 + \frac{\mathbf{c}}{\log N}.$$

代入（53），我们对任意 $t \in T$进行估计

$$F_T(x) \le |F(x,t)| + \frac{S_T(x)}{N^{1/2}} + \sum_{v:\, \frac{S_T(x)}{N^{1/2}} \le 2^{-v} \le S_T(x)} \mathcal{V}^r(x) \le |F(x,t)| + \frac{S_T(x)}{N^{1/2}} + \log N \cdot \mathcal{V}^r(x),$$

这表明

$$\|F_T\|_{L^2} \le \mathbf{C} \cdot \Big(A + \log N \cdot \|\mathcal{V}^r\|_{L^2} \Big), \qquad r = 2 + \frac{\mathbf{c}}{\log N},$$

因此，控制 r-变差算子，实质上便可推出（55）。

V^r 的相关估计在许多情况下都源于以下不等式；该不等式经典地用作鞅理论中的收敛结果 [26]；关于此不等式的讨论，见 [21]；更特殊的例子则见 [17] 或 [32]。

命题4.2 (Lépingle不等式，特殊情形)。—— 以下估计在条件期望情形下成立 (49)：

$$\|\mathcal{V}^r\|_{L^2(\mathbb{R})} \le \mathbf{C} \cdot \frac{r}{r-2} \cdot A.$$

命题 4.2 同样可以推广到卷积算子的情形（52）：通过结合一种平方函数论证

$$\begin{aligned}\mathcal{V}^r(\chi_k * f : k) &\le \mathcal{V}^r(\mathbb{E}_k f : k) + \mathcal{V}^r(\chi_k * f - \mathbb{E}_k f : k)\\ &\le \mathcal{V}^r(\mathbb{E}_k f : k) + 2 \cdot \Big(\sum_k |\chi_k * f - \mathbb{E}_k f|^2\Big)^{1/2}\end{aligned} \tag{58}$$

利用上文引入的估计式（51），可以运用正交性技巧和命题 4.2 来界定（58）中的两项。上文定义了离散时间变差

$$\mathcal{V}^r(f_k : k)(x) := \sup \Big(\sum_i |f_{k_i}(x) - f_{k_{i-1}}(x)|^r\Big)^{1/r}$$

其中，上确界取遍所有有限子序列 $\{k_i\}$。

W尽管 $\mathbf{V}^r$ 算子比相关的极大函数更为精细 ，

$$F_T(x) \le |F(x,t)| + \mathcal{V}^r(x)$$

对于任意 $t \in T$，它们本质上具有偶数阶的强度，具体来说，我们有如下启发式结论：

启发式结论 4.3。——无论是（49）还是（52），当以下两者都满足时，$\mathcal{V}^r$ 很难变大：F_I 以及平方函数

$$S_I(x) := \Big(\sum_{k:2^k \in I} |F(x,2^k) - F(x,2^{k+1})|^2\Big)^{1/2}$$

较小：$\mathcal{V}^r$“ $\approx$”$\frac{r}{r-2}\cdot(F_I + S_I)$“ $\approx$”$\frac{r}{r-2}\cdot F_I$，$\frac{r}{r-2}\cdot S_I$.[3]

最后，重要的是，鉴于我们以向量值视角研究

$$\{f_{\theta_1}, \dots, f_{\theta_N}\},$$

参见（34），我们注意到，正如极大函数和平方式函数一样，$\mathcal{V}^r$算子在向量值情形下也具有良好的相互作用：对于序列空间值函数$\vec{F} = (F_1, F_2, \dots)$

$$\mathcal{V}^r_{\vec{F}}(x) := \sup \Big(\sum_i \|F_n(x,t_i) - F_n(x,t_{i-1})\|^r_{\ell^2_n}\Big)^{1/r} \le \|\mathcal{V}^r_{F_n}(x)\|_{\ell^2_n}, \tag{59}$$

根据序列空间上的闵可夫斯基不等式（如 $r > 2$），其中上确界取遍$\{t_i\}$的有限递增子序列。

考虑到本节内容，我们开始看出 Bourgain 是如何展开论证的。

(3)在两种情形（49）或（52）下，极大函数与平方函数之间的密切联系是经典结果；例如参见 [**34**]。在情形（$\mathcal{V}^r, F_I, S_I$（49）中，它们之间的关系通过以下 good-λ不等式给出：

$$|\{\mathcal{V}^r > \mathbf{C}\lambda, F_I, S_I \le \gamma\lambda\}| \le \mathbf{C} \cdot \big(\frac{r}{r-2}\big)^2 \cdot \gamma^2 \cdot |\{\mathcal{V}^r > \lambda\}|, \quad r > 2;$$

见 [**23**]。

5. 论证逐渐成形

Bourgain 的任务是证明（37），这里我们只考虑 2^k 相对于 $|I| = \mathbf{C}$ 非常大的情形。根据单调收敛，我们可以将范围限制在有限多个尺度 $k \in T \subset \mathbb{N} \cap [\mathbf{C}\ \infty)$。我们聚焦于单个区间的情形。

受启发式分析的指导，我们令 $x_I \in I$ 为一个稍后确定的点，并试图界定

$$\| \sup_k | \sum_{n \le N} e(\theta_n x) \cdot \chi_k * f_{\theta_n}(x_I) | \|_{L^2_x(I)}.$$

如上所述——我们实际上被迫采用熵方法。具体而言，我们令

$$X(x_I) := \{ \chi_k * \vec{f_\Theta}(x_I) := \big(\chi_k * f_{\theta_1}(x_I), \dots, \chi_k * f_{\theta_N}(x_I) \big) : k \} \tag{60}$$

$\vec{}$，并令 N_λ 表示高度 λ 处相对于序列空间范数 $\ell^2([N])$ 的适当跳跃计数函数，

$$\vec{N_\lambda}(x) := \text{上确界} \Big\{ K : \text{存在一个时间序列 } \mathbf{C} \le k_0 < k_1 < \cdots < k_K \\ : \\ \| \chi_{k_i} * f_{\theta_n}(x) - \chi_{k_{i-1}} * f_{\theta_n}(x) \|_{\ell^2_{n \in [N]}} > \lambda \Big\}.$$

通过上述论证，我们可以对其进行估计

$$\begin{aligned} &\| \sup_k | \sum_{n \le N} e(\theta_n x) \cdot \chi_k * f_{\theta_n}(x_I) | \|_{L^2_x(I)} \\ &\le |\Xi_{k_0} * f(x_I)| \cdot |I|^{1/2} + \sum_v \| \max_{k \in \Lambda_v(x_I)} | \sum_{n \le N} e(\theta_n x) \cdot \big(\chi_k - \chi_{\varrho(k)} \big) * f_{\theta_n}(x_I) | \|_{L^2_x(I)} \end{aligned} \tag{61}$$

对任意 $k_0 \ge \mathbf{C}$，参见（30）。第一项性质较简单，因此我们暂且略去；对于每个单独的 v，我们可以估计

$$\begin{aligned} &\| \max_{k \in \Lambda_v(x_I)} | \sum_{n \le N} e(\theta_n x) \cdot \big(\chi_k - \chi_{\varrho(k)} \big) * f_{\theta_n}(x_I) | \|_{L^2_x(I)} \\ &\le \| \Big(\sum_{k \in \Lambda_v(x_I)} | \sum_{n \le N} e(\theta_n x) \cdot \big(\chi_k - \chi_{\varrho(k)} \big) * f_{\theta_n}(x_I) |^2 \Big)^{1/2} \|_{L^2_x(I)} \\ &\le \Big(\sum_{k \in \Lambda_v(x_I)} \| \sum_{n \le N} e(\theta_n x) \cdot \big(\chi_k - \chi_{\varrho(k)} \big) * f_{\theta_n}(x_I) \|^2_{L^2_x(I)} \Big)^{1/2} \\ &\le \mathbf{C} \cdot 2^{-v} \cdot \vec{N}_{2^{-v}}(x)^{1/2} \cdot |I|^{1/2} \end{aligned} \tag{62}$$

按照（46）中的论证，应用（34）来估计

$$\| \sum_{n \le N} e(\theta_n x) \cdot \big(\chi_k - \chi_{\varrho(k)} \big) * f_{\theta_n}(x_I) \|_{L^2_x(I)} \le \mathbf{C} \cdot 2^{-v} \cdot |I|^{1/2}$$

对 $k \in \Lambda_v(x_I)$ 一致成立。上文中的 $\{\Lambda_v(x_I) : v\}$ 是满足以下性质的极小时间集合：

$$X(x_I) \subset \bigcup_{k \in \Lambda_v(x_I)} B_{\ell^2([N])}\big(\chi_k * \vec{f_\Theta}(x_I), 2^{-v} \big),$$

其中，$B_{\ell^2([N])}(\vec{v},\ r)$ 是以 $\vec{v}\in\ell^2([N])$ 为中心、半径为 r 的球，度量采用序列空间范数 $\ell^2([N])$；父函数 ϱ 同上。

问题在于可能发生爆炸

$$\vec{N}_{2^{-v}}(x_I)\to\infty \text{ as } v\to\infty,$$

而且没有先验方法排除这个障碍；若有的话，（55）中就不会出现对数损失。Bourgain 的巧妙洞见使他得以突破这一抽象问题，而这洞见正是柯西–施瓦茨不等式：我们给出如下界

$$|\sum_{n\le N} e(\theta_n x)\cdot\big(\chi_k-\chi_{\varrho(k)}\big)*f_{\theta_n}(x_I)|\le N^{1/2}\cdot\Big(\sum_{n\le N}|\big(\chi_k-\chi_{\varrho(k)}\big)*f_{\theta_n}(x_I)|^2\Big)^{1/2}\le \mathbf{C}\cdot N^{1/2}\cdot 2^{-v}$$

对 $k\in\Lambda_v(x_I)$ 一致成立，从而得到一个容易的界

$$\|\max_{k\in\Lambda_v(x_I)}|\sum_{n\le N} e(\theta_n x)\cdot\big(\chi_k-\chi_{\varrho(k)}\big)*f_{\theta_n}(x_I)|\|_{L^2_x(I)}\le \mathbf{C}\cdot 2^{-v}\cdot N^{1/2}\cdot|I|^{1/2}.$$

总而言之，Bourgain 得到了如下界

$$\|\max_{k\in\Lambda_v(x_I)}|\sum_{n\le N} e(\theta_n x)\cdot\big(\chi_k-\chi_{\varrho(k)}\big)*f_{\theta_n}(x_I)|\|_{L^2_x(I)}\le \mathbf{C}\cdot 2^{-v}\cdot\min\{\vec{N}_{2^{-v}}(x_I)^{1/2},N^{1/2}\}\cdot|I|^{1/2},$$

见（62），他按照（57）巧妙地对其进行了插值，

$$2^{-v}\cdot\min\{\vec{N}_{2^{-v}}(x_I)^{1/2},N^{1/2}\}\le 2^{-v}\cdot\vec{N}_{2^{-v}}(x_I)^{1/r}\cdot N^{1/2-1/r}\le N^{1/2-1/r}\cdot\mathcal{V}^r_{\vec{f}_\Theta}(x_I)\le\mathbf{C}\cdot\mathcal{V}^r_{\vec{f}_\Theta}(x_I)\qquad r=2+\frac{\mathbf{c}}{\log N}$$

参见（59）和（60），从而得到右侧一个与 v无关的项。代入这个界，并按照启发式分析（53）中的方法进行推导，

$$(61)\le\mathbf{C}\cdot\sum_{v:2^{-v}\le\mathcal{F}_\Theta(x_I)}2^{-v}\cdot\min\{\vec{N}_{2^{-v}}(x_I)^{1/2},N^{1/2}\}\cdot|I|^{1/2}$$

$$\le\mathbf{C}\sum_{v:2^{-v}\le\mathcal{F}_\Theta(x_I)/N^{1/2}}2^{-v}\cdot N^{1/2}\cdot|I|^{1/2}+\mathbf{C}\sum_{v:\mathcal{F}_\Theta(x_I)/N^{1/2}\le 2^{-v}\le 2\cdot\mathcal{F}_\Theta(x_I)}\mathcal{V}^r_{\vec{f}_\Theta}(x_I)\cdot|I|^{1/2}$$

(63)

$$\le\mathbf{C}\cdot\mathcal{F}_\Theta(x_I)\cdot|I|^{1/2}+\mathbf{C}\cdot\log N\cdot\mathcal{V}^r_{\vec{f}_\Theta}(x_I)\cdot|I|^{1/2},\qquad r=2+\frac{\mathbf{c}}{\log N}.$$

最后，仔细选择 x_I 后，我们对其进行界定

$$\|\mathcal{M}_\Theta f\|_{L^2(I)}\le\mathbf{C}\cdot\|\mathcal{F}_\Theta\|_{L^2(I)}+\mathbf{C}\cdot\log N\cdot\|\mathcal{V}^r_{\vec{f}_\Theta}\|_{L^2(I)},$$

这表明，在尺度为C的 L^2-平均意义下，所有位置都满足以下不等式

$$\mathcal{M}_\Theta f\ \text{“}\le\text{”}\ \mathcal{F}_\Theta+\log N\cdot\mathcal{V}^r_{\vec{f}_\Theta},\qquad r=2+\frac{\mathbf{c}}{\log N}.$$

换言之，忽略对数误差后，向量值极大函数和向量值变差控制了 $\mathbf{M}_\Theta$。此外，对 $\{|I|=\mathbf{C}\}$ 求平方和，并考虑到

相关的、基于卷积的 Lépingle不等式版本，命题 4.2，可得如下界

$$\|\mathcal{M}_\Theta f\|_{L^2(\mathbb{R})} \le \mathbf{C}\cdot(1+\log N\cdot\frac{r}{r-2})\cdot\|f\|_{L^2(\mathbb{R})} \le \mathbf{C}\cdot(1+\log^2 N)\cdot\|f\|_{L^2(\mathbb{R})},$$

其满足（32）。

受这一直觉的启发，我们转而进行严格证明。

6. 命题2.3的证明：多频率极大不等式

受前述证明纲要的启发，我们将尝试证明以下估计：

$$\|\mathcal{M}_\Theta f\|_{L^2(\mathbb{R})} \le \mathbf{C}\cdot\log^2 N\cdot\|f\|_{L^2(\mathbb{R})}.$$

因此，我们只考虑尺度 $k \ge \mathbf{C}\log^2 N$，并用平方函数论证处理其余情形

$$\begin{aligned}\|\sup_{\mathbf{C}\le k\le \mathbf{C}\cdot\log^2 N}|\Xi_k f|\|_{L^2(\mathbb{R})} &\le \|\Big(\sum_{\mathbf{C}\le k\le \mathbf{C}\cdot\log^2 N}|\Xi_k f|^2\Big)^{1/2}\|_{L^2(\mathbb{R})}\\ &\le \mathbf{C}\cdot\log N\cdot\sup_k\|\Xi_k f\|_{L^2(\mathbb{R})} = \mathbf{C}\cdot\log N\cdot\sup_k\|\widehat{\Xi_k f}\|_{L^2(\mathbb{R})}\\ &\le \mathbf{C}\cdot\log N\cdot\sup_k\|\widehat{\Xi_k}\|_{L^\infty(\mathbb{R})}\cdot\|\hat f\|_{L^2(\mathbb{R})} \le \mathbf{C}\cdot\log N\cdot\|f\|_{L^2(\mathbb{R})},\end{aligned}$$

as

$$\sup_k\sup_\xi|\widehat{\Xi_k}(\xi)|\le 1,$$

见式（31）。因此，我们重新定义 $\mathcal{F}_\Theta = \mathcal{F}_{\Theta,\ \log^2 N}$见式（35），并注意到其继承的光滑性

$$\begin{aligned}(64)\quad |\mathcal{F}_\Theta(x)-\mathcal{F}_\Theta(y)| &\le \mathbf{C}\cdot\sum_{|I|=2^k\ge\log^2 N}\sum_{n\le N}|\chi_k * f_{\theta_n}(x)-\chi_k * f_{\theta_n}(y)|\\ &\le \mathbf{C}\cdot N\cdot\sum_{k\ge\log^2 N}(|x-y|\cdot 2^{-k})\cdot M_{\mathrm{HL}}f(x) \le \mathbf{C}\cdot|x-y|\cdot N^{-100}\cdot M_{\mathrm{HL}}f(x),\end{aligned}$$

w（即），这意味着 F_Θ 在尺度 $|I|=\mathbf{C}$上非常光滑。上文中，我们使用了以下界

$$|\partial\chi_k(x)| \le \mathbf{C}\cdot 2^{-k}\cdot 2^{-k}\cdot(1+2^{-k}\cdot|x|)^{-100}.$$

这种对尺度的剔除，使我们在严格论证这一启发式结论（33）时，不必像 Bourgain 那样谨慎：Bourgain 使用了一种所谓的最佳常数论证，而我们只需使用下面这个估计，它对于相对于

尺度 $k \geq \mathbf{C}\cdot \log^2 N$:

$$\begin{aligned}\sum_{n\leq N} e(\theta_n x)\cdot \chi_k * f_{\theta_n}(x) &= \sum_{n\leq N} e(\theta_n x)\cdot \chi_k * f_{\theta_n}(x_I) + O\Big(\frac{N\cdot |I|}{2^k}\cdot M_{\mathrm{HL}} f(x)\Big)\\ &= \sum_{n\leq N} e(\theta_n x)\cdot \chi_k * f_{\theta_n}(x_I) + O\Big(2^{-k/2}\cdot M_{\mathrm{HL}} f(x)\Big)\end{aligned}$$

对任意 $x_I \in I$，当然只要 $|I| \leq N^{\mathbf{C}}$。

特别地，对于任意 $x \in I$，当 $|I| = \mathbf{C}$ 时，我们可以给出上界

$$\sup_{k\geq \mathbf{C}\cdot\log^2 N} |\Xi_k f(x)| \leq \sup_{k\geq \mathbf{C}\cdot\log^2 N} \Big|\sum_{n\leq N} e(\theta_n x)\cdot \chi_k * f_n(x_I)\Big| + O\Big(\sum_{k\geq \mathbf{C}\cdot\log^2 N} 2^{-k/2}\cdot M_{\mathrm{HL}} f(x)\Big),$$

因此，对每个 I，我们可以给出上界

$$\begin{aligned}\|\sup_{k\geq \mathbf{C}\cdot\log^2 N} |\Xi_k f(x)|\|_{L^2(I)} \leq \mathbf{C}\cdot \min_{x_I\in I} \|\sup_{k\geq \mathbf{C}\cdot\log^2 N} \Big|\sum_{n\leq N} e(\theta_n x)\cdot \chi_k * f_n(x_I)\Big|\|_{L^2(I)}\\ + \mathbf{C}\cdot N^{-100}\cdot \|M_{\mathrm{HL}} f\|_{L^2(I)}\end{aligned}$$

（例如）。暂时略去含有 M_{HL} 的项，因为它并不重要，我们来考察第一项

$$\|\sup_{k\geq \mathbf{C}\cdot\log^2 N} \Big|\sum_{n\leq N} e(\theta_n x)\cdot \chi_k * f_n(x_I)\Big|\|_{L^2(I)},$$

我们使用熵方法对其进行估计，见（63），

$$\begin{aligned}&\|\sup_{k\geq \mathbf{C}\cdot\log^2 N} \Big|\sum_{n\leq N} e(\theta_n x)\cdot \chi_k * f_n(x_I)\Big|\|_{L^2(I)}\\ &\quad\leq \mathbf{C}\cdot \Big(\mathcal{F}_\Theta(x_I)\cdot |I|^{1/2} + \log N\cdot \mathcal{V}^r_{\vec{f_\Theta}}(x_I)\cdot |I|^{1/2}\Big),\\ &\quad\quad\leq \mathbf{C}\cdot \Big(\|\mathcal{F}_\Theta\|_{L^2(I)} + \log N\cdot \|\mathcal{V}^r_{\vec{f_\Theta}}\|_{L^2(I)} + N^{-100}\cdot \|M_{\mathrm{HL}} f\|_{L^2(I)}\Big),\end{aligned}$$

选择 x_I 使其在 I 上使 $\mathbf{V}^r_{\vec{f_\Theta}}$ 最小，并利用光滑性

$$\mathcal{F}_\Theta(x_I) = \mathcal{F}_\Theta(x) + O(N^{-100}\cdot M_{\mathrm{HL}} f(x))$$

来界定

$$\mathcal{F}_\Theta(x_I)\cdot |I|^{1/2} = \|\mathcal{F}_\Theta\|_{L^2(I)} + O\Big(N^{-100}\cdot \|M_{\mathrm{HL}} f\|_{L^2(I)}\Big).$$

特别地，我们得到了有界性

$$\|\sup_{k\geq \mathbf{C}\cdot\log^2 N} |\Xi_k f|\|_{L^2(I)} \leq \mathbf{C}\cdot \Big(\|\mathcal{F}_\Theta\|_{L^2(I)} + \log N\cdot \|\mathcal{V}^r_{\vec{f_\Theta}}\|_{L^2(I)} + N^{-100}\cdot \|M_{\mathrm{HL}} f\|_{L^2(I)}\Big)$$

其中 $r = 2 + \mathbf{c}\log N$，因此对 $|I| = \mathbf{C}$ 求平方和，最终得到如下界

$$\begin{aligned}\|\mathcal{M}_\Theta f\|_{L^2(\mathbb{R})} &\le \Big(\sum_{|I|=\mathbf{C}} \| \sup_{k\ge \mathbf{C}\cdot\log^2 N} |\Xi_k f|\|_{L^2(I)}^2\Big)^{1/2} + \mathbf{C}\cdot\log N\cdot\|f\|_{L^2(\mathbb{R})}\\ &\le \mathbf{C}\cdot\Big(\sum_{|I|=\mathbf{C}} \|\mathcal{F}_\Theta\|_{L^2(I)}^2\Big)^{1/2} + \mathbf{C}\cdot\log N\cdot\Big(\sum_{|I|=\mathbf{C}} \|\mathcal{V}^r_{\vec{f}_\Theta}\|_{L^2(I)}^2\Big)^{1/2}\\ &\qquad + \mathbf{C}\cdot N^{-100}\cdot\Big(\sum_{|I|=\mathbf{C}} \|M_{\mathrm{HL}} f\|_{L^2(I)}^2\Big)^{1/2} + \mathbf{C}\cdot\log N\cdot\|f\|_{L^2(\mathbb{R})}\\ &\le \mathbf{C}\cdot\Big(\|\mathcal{F}_\Theta\|_{L^2(\mathbb{R})} + \log N\cdot\|\mathcal{V}^r_{\vec{f}_\Theta}\|_{L^2(\mathbb{R})} + N^{-100}\cdot\|M_{\mathrm{HL}} f\|_{L^2(\mathbb{R})} + \log N\cdot\|f\|_{L^2(\mathbb{R})}\Big)\\ &\le \mathbf{C}\cdot\log^2 N\cdot\|f\|_{L^2(\mathbb{R})},\end{aligned}$$

完成证明。

7. 当代研究

自 Bourgain 的研究以来，Mariusz Mirek、Eli Stein 及其合作者（[27]； [28]； [29]； [30]； [31]）接续并大大推进了多项式轨道上的遍历平均逐点收敛这一课题，其研究建立在 [19]的突破性工作之上。当前的研究现状由 [30]确立：

定理 7.1。— 设(X,μ) 是一个 σ-有限测度空间， $\tau:\ X\to X$ 是一个保测变换，而 $P\in\mathbb{Z}[\cdot]$ 是一个整系数多项式。那么对每个 $1<p<\infty$、 $r>2$

$$\begin{aligned}&\|\mathcal{V}^r\Big(\frac{1}{N}\sum_{n=1}^N \tau^{P(n)} f : N\Big)\|_{L^p(X)} + \sup_{\lambda>0}\|\lambda\cdot N_\lambda\Big(\frac{1}{N}\sum_{n=1}^N \tau^{P(n)} f : N\Big)^{1/2}\|_{L^p(X)}\\ &\qquad\le \mathbf{C}_p\cdot(1+\frac{r}{r-2})\cdot\|f\|_{L^p(X)}.\end{aligned}$$

换言之，从定量的角度来看，抽象平均值的收敛速率

$$\frac{1}{N}\sum_{n=1}^N \tau^{P(n)} f$$

正是我们那些熵例子的特征！

这一论证的关键在于将 $\mathbb{Q}\cap[0,1]$ 组合划分为所谓的有理数 Ionescu–Wainger 穷竭：人们用.....替代

$$\mathcal{R}_s \longrightarrow \mathcal{U}_s,$$

见（28），其中 $\{\mathcal{U}_s:\ s\}$ 构成 $\mathbb{Q}\cap[0,1]$ 的不交划分，它捕捉了 $\mathcal{R}_s$的许多解析性质，即

$$\sup_{A/Q\in\mathcal{U}_s} |S(A/Q)| \le \mathbf{C}_\epsilon\cdot 2^{(\epsilon-1/d)s}, \tag{65}$$

但 adm　它具有更优的算术统计性质，从而可以进行近似

$$K'_k \quad \text{“}=\text{”} \quad \sum_{s\le \mathbf{c}\cdot k}\ \sum_{A/Q\in\mathcal{U}_s} S(A/Q)\cdot \mathrm{Mod}_{A/Q}(\chi_{(d-\mathbf{c})k} * \phi'_{dk}) =: \sum_{s\le \mathbf{c}\cdot k} L_{k,s}$$

在 $\ell^p(\mathbb{Z})$ 中也成立。

尽管 Bourgain 的熵论证在一般情况下对 ℓ^p，$p\neq 2$ 的效果较差，但通过应用拉德马赫–门绍夫不等式，并按照 [5]，中的方法论证，可以得到例如如下估计

$$\|\sup_k |L_{k,s} * f|\|_{\ell^p(\mathbb{Z})} \le \mathbf{C}_\epsilon \cdot 2^{\epsilon s}\cdot 2^{-\mathbf{c}_{p,d}s}\cdot \|f\|_{\ell^p(\mathbb{Z})}, \qquad 1<p<\infty,\ \mathbf{c}_{p,d}<1/d,$$

跳跃计数表述也类似。频率数量上的损失在 s 中是次指数级的，这与 Bourgain 极大函数在 ℓ^2 上的情形相同；而增益

$$2^{-\mathbf{c}_{p,d}s}$$

可通过对（65）作适当插值而得。

相较于关于 $\sup_k |L'_{k,s} * f|$ 的最精确估计，这一定量改进

$$\|\sup_k |L'_{k,s} * f|\|_{\ell^p(\mathbb{Z})} \le \mathbf{C}_{\epsilon,p}\cdot 2^{(\epsilon+1)s}\cdot 2^{-\mathbf{c}_{p,d}s}\cdot \|f\|_{\ell^p(\mathbb{Z})}, \quad 1<p<\infty,\ \mathbf{c}_{p,d}<1/d,$$

体现了这些论证的灵活性；事实上，它们还可扩展到处理 r-变差和跳跃计数算子的情形。

参考文献

[1] A. Bellow. 两个问题. 《数学讲义》945, Springer-Verlag, 柏林, 第 429-431 页。[2] A. Bellow; V. Losert. 遍历理论中密度为零的序列, Contemp. Math. 26 (1984), 49-60.[3] G. Birkhoff. 遍历定理的证明。Proc Natl Acad Sci USA 17 (12): 656-660 (1931)[4] J. Bourgain. 与凸体相关的高维极大函数。Amer. J. Math. 108 (1986), no. 6, 1467–1476.[5] J. Bourgain. 关于正整数某些子集上的极大遍历定理。Israel J. Math. 61 (1988), 39-72.[6] J. Bourgain. 关于 L^p 算术集合上的逐点遍历定理。Israel J. Math. 61 (1988), no. 1, 73-84.[7] J. Bourgain. 几乎必然收敛与有界熵。Israel J. Math. 63 (1988), no. 1, 79–97.[8] J. Bourgain. 有界正交系与 $\Lambda(p)$-集问题。Acta Math. 162 (1989), no. 3-4, 227–245.

[9] J. Bourgain. *Pointwise ergodic theorems for arithmetic sets.* Inst. Hautes Études Sci. Publ. Math. (69):5-45, 1989. 附有作者 Harry Furstenberg、Yitzhak Katznelson 和 Donald S. Ornstein 撰写的附录。[10] J. Bourgain. 双重再现与几乎必然收敛。J. Reine Angew. Math. 404 (1990), 140–161.[11] Z. Buczolich; D. Mauldin. 平方数上的发散遍历平均背后的概念。遍历理论及相关领域，美国数学会，当代数学第 430 卷 (2007) 41-56[12] A. Calderón. 遍历理论与平移不变算子。美国国家科学院院刊 59 (1968), 349-353[13] C. Demeter. 遍历双线性希尔伯特变换的逐点收敛。Illi- nois J. Math. 51 (2007), no. 4, 1123–1158.[14] C. Demeter; M. Lacey; T. Tao; C. Thiele. 打破返回时间定理中的对偶性。Duke Math. J. 143 (2008), no. 2, 281–355.[15] C. Demeter; T. Tao; C. Thiele. 极大多线性算子。Trans. Amer. Math. Soc. 360 (2008), no. 9, 4989–5042.[16] H. Furstenberg. Durham会议，1982年6月[17] S. Guo; J. Roos; P.-L. Yung. 与Carleson定理相关的振荡积分的尖锐变差范数估计。Anal. PDE 13 (2020), no. 5, 1457–1500.[18] L.-K. Hua. 数论导引。Springer-Verlag，柏林—纽约，1982。Peter Shiu 从中文译出。[19] A. Ionescu; S. Wainger. L^p 离散奇异Radon变换的有界性。J. Amer. Math. Soc. 19, (2005), no. 2, 357—383.[20] R. Jones; R. Kaufman; J. Rosenblatt; M. Wierdl. 遍历理论中的振荡。Ergodic Theory Dynam. Systems 18 (1998), no. 4, 889-935.[21] R. Jones; A. Seeger; J. Wright. 调和分析中的强变分不等式与跳跃不等式。Trans. Amer. Math. Soc. 360 (2008), no. 12, 6711-6742.[22] R. Jones; G. Wang. Féjer核与Poisson核的变差不等式。Trans. Amer. Math. Soc., 356 (2004), 4493-4518.[23] B. Krause. 调和分析中的离散对应物：Bourgain、Stein 及其后续发展。即将作为数学研究生丛书第224卷出版。美国数学会，罗得岛州普罗维登斯，2023[24] M. Lacey. 双线性极大函数映射到 L^p，对于 $\frac{2}{3} < p \leq 1$。Ann. of Math. (2) 151 (2000), no. 1, 35–57.[25] P. LaVictoire. 普遍的 L^1-坏算术序列。J. Anal. Math. 113 (2011), 241–263.[26] D. Lépingle. 半鞅的 p 阶变差。Z.F.W. 36, 1976, 295-316.[27] M. Mirek. 离散Radon变换的平方函数估计。Anal. PDE 11 (2018), no. 3, 583–608.

[28] M. Mirek; E. Stein; B. Trojan. $\ell^p(\mathbb{Z}^d)$-离散 *Radon* 型算子的估计：极大函数与向量值估计。J. Funct. Anal. 277 (2019), no. 8, 2471–2521.[29] M. Mirek; E. Stein; B. Trojan. $\ell^p(\mathbb{Z}^d)$-离散 *Radon* 型算子的估计：变分估计。Invent. Math. 209 (2017), no. 3, 665–748.[30] M. Mirek; E. Stein; P. Zorin-Kranich.整数上的 *Radon* 型平移不变算子的跳跃不等式 $\mathbb{Z}^d$。Adv. Math. 365 (2020), 107065[31] M. Mirek; B. Trojan.高维离散极大函数及其在遍历理论中的应用。Amer. J. Math. 138 (2016), no. 6, 1495–1532.[32] R. Oberlin; A. Seeger; T. Tao; C. Thiele; J. Wright. 变差范数 *Carleson* 定理。J. Eur. Math. Soc. (JEMS) 14 (2012), no. 2, 421–464.[33] J. Qian. 部分和过程与经验过程的 p-变差。Ann. of Prob. Theory, 77 (1988), 1370-1383.[34] E. Stein. 调和分析：实变量方法、正交性与振荡积分。普林斯顿数学丛书第43卷。调和分析专著 III。普林斯顿大学出版社，普林斯顿，新泽西州，1993.[35] T. Tao. 探究 *Jean Bourgain* 的工具箱。Bull. Amer. Math. Soc. (N.S.) 58 (2021), no. 2, 155–171.[36]J.-P. Thouvenot. （根据 *Jean Bourgain*）某些整数子序列上的遍历平均几乎必然收敛布尔巴基研讨会，1989/90年卷。Astérisque 第189-190期 (1990), Exp. No. 719, 133–153.

Ben Krause 伦敦国王学院 数学系 斯特兰德街，伦敦 WC2R 2LS电子邮件：
ben.krause@kcl.ac.uk