韩国宫廷音乐中的齐夫–曼德布罗特标度：音乐中的普遍模式

Byeong chan Choi[1,+], **Junwon You**[1,+], **Myung Ock Kim**[2], **和 Jae-Hun Jung**[1,*]

[1]浦项科技大学，数学系，浦项，37673，韩国[2]韩国高等研究院（KIAS），首尔，02455，韩国[*]通讯作者
jung153@postech.ac.kr[+]这些作者对本研究贡献相同

arXiv:2602.14198v1 [stat.AP] 15 Feb 2026

摘要

齐夫定律最初发现于自然语言中，后来被推广为齐夫–曼德布罗特定律，它描述了齐夫元素的频率与其排名之间的幂律关系。由于这一规律的语义特征，它也在音乐数据中被观察到。然而，此类研究大多聚焦于西方音乐，其在非西方音乐中的适用性仍未得到充分研究。我们分析了43首韩国宫廷音乐乐曲，即正乐，这些乐曲跨越数个世纪，以韩国传统音乐记谱井间谱记录。这些乐曲被转写为西方五线谱，并从中提取音高、时值等音乐数据。我们将音高、时值及其配对组合用作齐夫单位，发现韩国音乐也高度符合齐夫–曼德布罗特定律，尤其是音高–时值配对单位。韩国音乐在漫长岁月中集体演变，逐渐平滑了个体差异，并形成了人们普遍能够理解的形式。这种集体演变似乎在塑造促成齐夫–曼德布罗特定律成立的特征方面发挥了重要作用。我们的发现为音乐数据中的齐夫–曼德布罗特标度具有跨文化普遍性提供了更多证据。我们进一步表明，两个相互独立的齐夫数据集的联合分布遵循齐夫–曼德布罗特定律；从这个意义上说，我们的结果不仅拓展了齐夫定律，也加深了我们对组合与交互作用下标度律行为的理解，为基于排名的统计规律提供了更统一的视角。

1 引言

最初在自然语言中发现的齐夫定律，描述了频数f、$f \in \mathbb{Z}_{>0}$与排名r、$r \in \mathbb{Z}_{>0}$之间的幂律关系；这些元素在数据集中称为齐夫元素[1,2]，其形式为

$$f(r) \sim \frac{1}{r^s} \tag{1}$$

其中s、$s \in \mathbb{R}_{>0}$为指数，即对数尺度下f的幂律衰减斜率。最初促成齐夫定律提出的发现表明，幂律衰减的斜率约为-1，即$s \sim 1$。斜率值s与数据集的语义结构有关。斜率$s \approx 1$通常被视为交流中平衡冗余与多样性的“效率”指标。在自然语言中，$s \approx 1$与认知或语义结构、最优信息传递以及交流效率相关联。因此，类比而言，如果音乐、生态学、遗传学等领域的类似数据集也呈现出相似的斜率值s，这可能意味着它同样反映了生成该数据集的系统的某种“语义”或“结构”属性[1–4]。正如齐夫定律所示，排名靠前的元素，其出现频数会随着排名上升而按代数规律衰减。因此，数据集中出现频数最高和最低的元素之间存在巨大差异。这一模式最初是在人类行为和心理学的背景下得到解释的，尤其是从交流效率和最省力原则的角度[1,2]。后续研究表明，齐夫定律并不限于自然语言数据集，在广泛的实证数据中也能观察到。例子包括城市人口分布[5]，、网站访问频率[6]，以及生态学中的物种丰度[7]。齐夫定律无处不在这一现象固然引人注目，但其广泛出现的根本原因仍未得到充分解释。

原始的齐夫定律假设系统理想且无限大，忽略了有限尺寸效应。在现实数据集中，出现频率最高的元素往往偏离纯幂律分布。为解释这类偏差，齐夫定律被推广为齐夫–曼德布罗特定律[3,8,9]，，其形式如下：

$$f(r) \sim \frac{1}{(r+q)^s}, \tag{2}$$

其中 r 为排名，s 为齐夫定律中的指数，q，$q \in \mathbb{R}_{\geq 0}$ 是一个额外参数，称为平移参数。它通过平移排名来考虑有限尺寸效应和高频偏差。研究表明，自然语言数据以及后来发现的其他复杂系统（如音乐、城市规模和收入分布）都系统性地偏离了理想的齐夫定律[3,8]。平移参数 q 能更准确地拟合经验频率分布。对于排名靠前的元素，当r较小时，q的影响显著，会使曲线变平。当r较大时，$r+q \approx r$，该定律便简化为标准的齐夫形式。这一修正能更好地拟合真实数据集。

鉴于其与语义结构的密切关联，原始齐夫定律已被用于刻画音乐数据集中的统计模式，尤其是西方音乐的统计模式。在西方音乐中，大量研究同时考察了原始齐夫定律及其齐夫–曼德布罗特扩展。例如，Voss 和 Clarke[10] 发现，音乐信号呈现出 $1/f$噪声特征，音高波动遵循反幂律分布。此处f是音乐信号中波动（如音高或振幅变化）的时间频率。这一发现将音乐定位于纯随机性与严格周期性之间，从而在可预测性（秩序）与惊奇度（随机性）之间形成平衡，在感知上令人愉悦。幂律斜率$s \approx 1$确保能量在各尺度上分布，形成分形结构，这意味着音乐既非完全随机，也非完全重复。Manaris 等人[11] 将齐夫–曼德布罗特定律应用于一个大型语料库，其中包含220首MIDI编码作品，涵盖巴洛克、古典主义、浪漫主义、十二音、爵士乐、摇滚乐以及算法生成音乐，表明这些指标能够以计算方式捕捉审美属性。本研究使用的指标包括音高、时值、旋律音程与和声协和。这些发现表明，遵循齐夫分布的指标会产生具有审美意义的音乐。这意味着齐夫定律与听众凭直觉用以理解和欣赏音乐的音乐语义之间存在关联。

将 Zipf 定律应用于音乐并非易事，因为这需要定义语义元素或齐夫单位，使其类似于自然语言中的词语。然而，在音乐中，这类语义元素尚未得到明确界定，不同研究也对各种定义进行了检验。Perotti 和 Billoni[12] 发现，将和弦和音符都视为齐夫单位时，会出现齐夫分布，这进一步强化了音乐与语言之间的类比。Zanette[13] 将语言语境的概念扩展至音乐，并表明音符使用的统计特征遵循西蒙模型——一种支撑齐夫定律的生成机制。此后，研究者还提出了进一步的改进。这里的西蒙模型[14] 是一种用于生成符合齐夫定律的分布的随机模型。该模型最初用于解释自然语言中的词频分布，但其应用范围更广。通过简单的随机过程，该模型便能轻松生成齐夫分布。在[13], 中，音乐音符或事件被视为序列中的元素，并运用西蒙模型解释音乐语境和音符使用如何产生类似齐夫定律的统计特征；重复某个音符的概率与该音符在音乐序列中出现的次数成正比，因此音符频率呈现幂律分布。这意味着，音乐可能与自然语言一样，具有“语义语境”：某些音符或动机在特定语境中会更频繁地重复出现。Serra-Peralta 等人[15]利用古典音乐的 Kunstderfuge 语料库[1]分析了古典音乐中的和声码字。他们使用了希普定律：$V(N) \sim KN^{\beta}$，其中 $V(N)$ 表示不同元素的数量（即词汇量），N 表示数据集中元素的总数，$K > 0$为常数，而 $\beta, 0 < \beta < 1$为指数。随后，他们展示了词汇量与齐夫定律之间的数学联系，并用 $\beta \approx {}^{1}_{s}$ 概括。也就是说，作者将不同的和声码字视为音乐词汇，并表明这些码字符合齐夫定律。Haro 等人[16] 分析了从 740 小时的语音、音乐和环境声音中提取的音色码字，并指出齐夫分布普遍存在于听觉信号中，其适用范围不仅涵盖自然语言，也包括音乐和环境声音。这一结果支持了这样一种观点：齐夫定律是结构化符号序列的普遍特征，涵盖音乐音色，而不仅仅是语言中的词序列。

尽管关于音乐中齐夫定律的研究十分广泛，但其应用主要集中于西方音乐。如上所述，齐夫定律揭示了一个重要含义：音乐的构建可能依赖于潜在的音乐语义，而听众会凭直觉运用一种统计框架来理解和欣赏乐曲。此前研究的实证证据支持这一观点，因此它与音乐学和美学分析密切相关。然而，这类证据在很大程度上仅限于西方音乐；类似模式是否也存在于其他音乐传统中，仍不明确，尤其是在音乐理论体系尚未确立或受到较少关注的数据集中。与拥有完善理论框架（如和声与对位法）的西方音乐不同，许多民族音乐传统缺乏此类形式化体系。

在此背景下，我们分析了一个大规模的韩国传统宫廷音乐数据集——正乐 ，以探究其统计分布，并评估其是否也遵循齐夫定律。与西方乐曲不同，韩国传统音乐并非依据统一的正式音乐理论创作。相反，每首乐曲似乎都有各自的作曲原则。与西方音乐相比，韩国音乐的结构规范性较低、变化更多，因此对其进行统计分析更具挑战性。

如上文所述，要应用齐夫定律，定义数据集的语义元素，即齐夫单位，非常重要。继先前的研究（如[10,12,13], ）之后，我们将音高、时值及其成对组合视为齐夫单位，用于

[1]https://www.kunstderfuge.com/

韩国传统音乐。韩国传统音乐本质上是异音的，并不包含和声这一概念。因此，在本研究中，基于和声结构的齐夫分析并不适用。

本研究分析了43首正乐乐曲，这些乐曲采用传统韩国音乐记谱法井间谱记谱，年代跨越数百年，最古老的曲目距今约有600年。尽管现存韩国传统音乐的语料库规模更大，但我们选择这些乐曲，是因为它们以书面形式保存下来，不同于许多口传作品。正乐主要在宫廷和贵族阶层中演奏，其记谱文献使准确的统计分析成为可能。正如第2节所述，我们将井间谱乐谱转换为西方五线谱，以便兼容标准计算音乐分析工具，随后进行了仔细的预处理和归一化。KIAS大型韩国音乐研究组将井间谱记谱转换为西方五线谱。在此过程中，音高和时值的转换依据不同的音乐惯例分别进行，因此我们对数据进行了归一化处理，将其合并为一个数据集以供分析[17]。

我们的分析表明，韩国传统宫廷音乐在所有齐夫单位上都遵循齐夫–曼德布罗特定律，包括音高、时值及其组合。尤其是音高与时值构成的配对单位，清楚地表明韩国音乐遵循与西方音乐相同的统计模式。此外，我们的结果显示，与西方音乐相比，韩国音乐对齐夫–曼德博分布的遵循程度甚至更高。这可以用韩国传统音乐的集体性和长期发展特征来解释：它是在集体创造而非个人创作的过程中演变而来的。

正如 Zipf[1,2], Mandelbrot[3,8], Voss 和 Clarke[10], Serra-Peralta 等人[15] 及其他学者所指出的，音乐中存在齐夫定律，这表明音乐序列围绕语义单位构建，使听众能够像理解和欣赏自然语言一样感知和欣赏音乐。我们发现，韩国传统音乐同样遵循齐夫定律，这说明其具有类似的内在语义结构。此外，它对齐夫定律的遵循程度更高，这表明随着时间推移，韩国音乐在集体作用下逐渐形成，变得更易于听众理解，反映出一种更具集体性、而非个人性的创作过程。换言之，数百年来，个体特异性效应逐渐被消磨，音乐形式因而受到集体文化共识的塑造。这类长期、大规模且自组织的过程，正是齐夫型分布通常出现的条件，自然语言便是如此。相比之下，单一作曲家的作品或昙花一现的音乐风格，可能缺乏形成稳定齐夫型模式所需的规模和多样性。Corral等人[18] 表明，与短篇文本或单一作者样本相比，齐夫定律在长期积累的集体文本中更为稳健，凸显了规模和文化积累的稳定作用。同样，Serra-Peralta 等人[15] 发现，相比孤立作品，音乐中的齐夫型标度在规模更大、更多样化的语料库中表现得更为明显。

我们进一步表明，两个独立齐夫型数据集的联合分布遵循齐夫–曼德博分布。这一结果揭示了标度律在组合与相互作用下的表现，并为基于排名的统计规律提供了更统一的视角。这一观察同样适用于本研究所考察的韩国音乐数据。具体而言，对于音高与时值的联合分布，其各边缘数据集均至少满足分段齐夫定律。由于自然界中的大多数数据都以两个或更多变量的联合分布形式出现，这一结果或许有助于解释齐夫–曼德博分布为何常见于实证数据中。

本文结构安排如下。第 2 节介绍本研究使用的数据集。具体而言，该数据集源自大规模井间谱乐谱，这些乐谱先被转换为西方五线谱记谱，随后经过预处理和归一化。本节还定义了分析中使用的齐夫单位。第 3 节介绍齐夫–曼德博定律，并讨论其参数的含义。本节还将说明估计这些参数所采用的方法。第 4 节对数据进行详细分析，并证明数据遵循齐夫–曼德博分布。第 5 节进行理论探讨，其中包括有关成对齐普夫数据的齐夫–曼德博定律的结果。最后，第 6 节总结全文，并讨论未来研究方向。

2 数据集与齐夫单位

韩国宫廷音乐以世宗大王（1397-1450）创制的井间谱记谱形式流传下来。在井间谱中，音高以称为律名的汉字书写。然而，井间谱 更根本的特点在于，它通过将声音置于一系列方格中，使我们能够辨识其时值。

我们分析了由KIAS大型韩国音乐研究组构建的韩国音乐数据。该数据包含45首以井间谱记谱的宫廷音乐曲目。所选曲目代表了目前在韩国国立国乐院演出的标准曲目库，相关井间谱 由国立国乐院于2015至2016年出版。由于井间谱 通常以分谱而非总谱的形式出版，这些数据被转换为220个MusicXML文件，每个文件代表一个独立的器乐声部。

本次分析中的所有曲目均为合奏作品，大多创作于15至19世纪。分析所涉及的乐器是韩国宫廷音乐中常见的乐器：大笒、觱篥、奚琴、伽倻琴、玄琴、牙箏 和人声。在这45首曲目中，仅收录了与民乐 七个乐章中的第1至第3乐章，而歌曲 原本

是以包含41首曲目的大型组曲形式演奏的，但本数据仅收录其中三首。研究中使用这些数据前，我们剔除了由两首曲目 보허자 和해령组成、受中国音乐影响的八个文件。中国音乐在很长一段时期内传入韩国；然而，其音乐风格与韩国音乐不同。自15世纪以来，人们创作了若干具有中国音乐风格的曲目。因此，音乐学家根据这些作品的中国风格特征，将其与韩国原创音乐区分开来。据此，本研究使用了数据中的43首曲目，共计212个musicXML文件。此外，在该数据集中，与民乐 和阳清道调被拆分为多个文件；拆分标准将在后文说明。本文所考察乐曲的完整列表见表3。

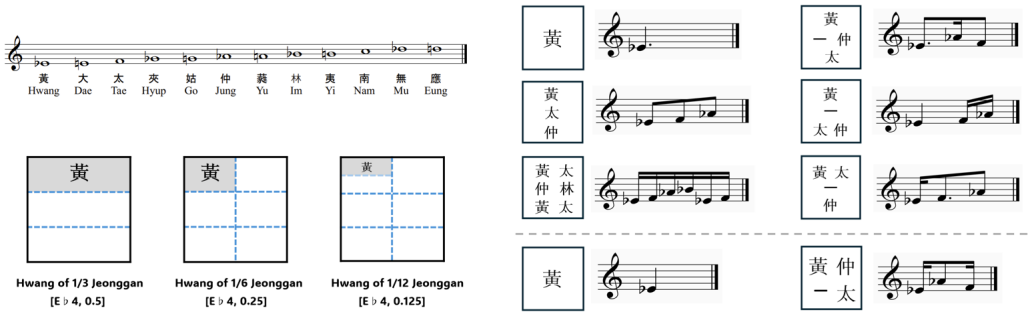


图1. 井间谱记谱的解读方式。韩国传统音名（律名）与相应的西方音高对应，其中黃（黃）对应 E♭4。时值通过井间的空间结构表示：一个方格表示一拍，可细分为两个或三个更小的单元，并可进一步递归细分。在这些细分格中，音高符号的位置和分组同时标示音高与时值，从而可以解读井间谱并将其转写为西方五线谱记谱。

称为律名的音名发源于古代中国，15世纪后出现在韩国的井间谱中。一个八度内有十二个律名，分别对应特定的西方音符（见图1）。在韩国音乐中，主音黃钟(黃)对应 E♭，十二个音高各有其名称。井间是井间谱记谱中的一个方格。音高符号在律名中所占的井间空间表示其时值。因此，律名被转换为音高，而井间面积被转换为时值。世宗时期的井间谱在20世纪改良为现代式井间谱，至今仍用于演奏。现代式井间谱包含更详细的音乐信息。一个井间表示一拍，可根据乐曲通常细分为三个或两个小拍。这些细分后的拍还可以继续反复二分（见图1）。这样一来，井间便能清楚地标示音高和时值，并可转换为五线谱记谱。

韩国传统音乐以支声复调著称，其中每个声部的旋律都比垂直关系更受重视。在一首乐曲中，各声部共享基本旋律线，但每个声部的旋律都是由独立的创作者创作而成。这一过程持续了数十年甚至数百年。因此，相较于各声部之间的垂直关系，每个声部的旋律与音乐创作过程的关联更为密切。基于这一原因，我们对乐曲中各声部的数据分别进行了分析。

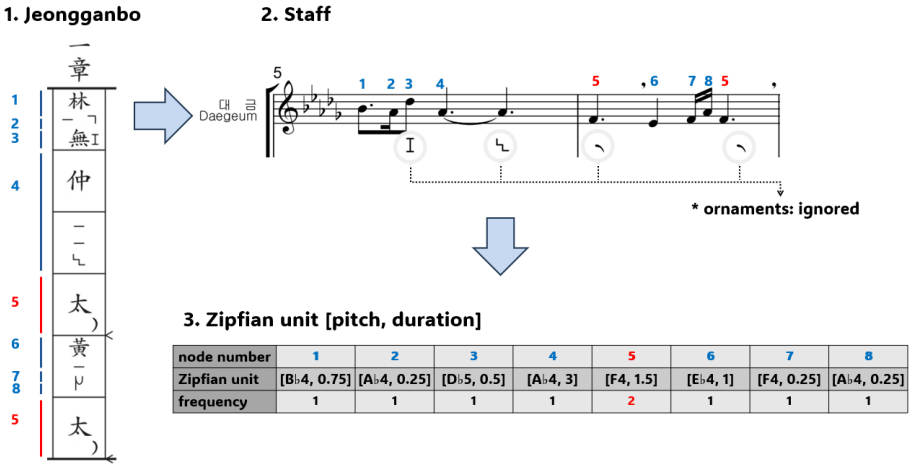
2.1 数据预处理与齐夫单位的定义

为考察齐夫定律对韩国传统音乐的适用性，首先将井间谱记谱转换为西方五线谱记谱，并以 MusicXML 格式存储[19]。随后，使用 music21 Python 库[20], 处理转换后的谱表，该库提供了强大的工具，可用于解析、分析和提取符号化音乐信息。

使用 music21.converter.parse() 函数加载并解析每个 MusicXML 文件。从解析后的数据中，我们提取了各个音乐事件的音高和时值信息。非音高元素（例如休止符、连音线和歌词）不纳入分析。仅保留有效音符：音符表示具有特定时值的单个音高事件。

韩国音乐的旋律由简单逐渐演变得细腻。乐曲形成的这一漫长过程对旋律的形成发挥了重要作用；因此，井间谱 除了记录音高和时值外，还记录了各种装饰音。然而，这种音乐仍包含即兴演奏，装饰音等音乐元素可能因演奏者而异。因此，律名 和时值是井间谱中旋律形成最基本的元素。（为考察）旋律的语义结构，尤其是从创作角度出发，本文仅考虑音高和时值。

本研究将传统的 井间谱 音名转换为西方音名（见图 2）。这种对应关系是参考韩国宫廷音乐的常见学术解读和演奏实践确定的。例如，



node number	1	2	3	4	5	6	7	8
Zipfian unit	[B♭4, 0.75]	[A♭4, 0.25]	[D♭5, 0.5]	[A♭4, 3]	[F4, 1.5]	[E♭4, 1]	[F4, 0.25]	[A♭4, 0.25]
frequency	1	1	1	1	2	1	1	1

图 2. 将井间谱转换为齐夫单位的流程概览。井间谱首先被转录为西方五线谱，以便使用 music21 库对韩国宫廷音乐进行计算分析。随后，提取（音高，时值）对作为齐夫单位，并计算其频数。忽略装饰音符号。

音符 黃钟 （黃）对应 E♭ 4，太 （太）对应 F4，以此类推。图 5 （见附录）总结了本研究采用的对应关系。

每个音符均表示为一个元组，其中包含音名（例如 E♭ 4、F4）和以井间单元为单位测量的节奏时值。我们将这些（音高，时值）元组定义为齐夫单位，以同时体现音乐结构的旋律和节奏维度。

为确保不同乐器演奏的曲目之间具有可比性，我们根据各乐器的典型音高范围进行了音高归一化。由于数据集包含由多种乐器演奏的分谱，例如大笒、觱篥、奚琴、伽倻琴、玄琴和牙箏，因此有必要进行音高归一化，使所有音乐事件的音区保持一致。

在韩国宫廷音乐中，“中央音域”这一概念早已得到认可。与此一致，一项关于韩国中世纪音乐的研究发现，音符出现频率在该中央八度音域内最高[21]。

尽管现代乐器能够产生比其中世纪对应乐器更宽的音高范围，音乐使用仍往往集中于每种乐器的特定音区[22]。因此，我们对各乐器最常出现的音高范围进行了归一化。具体而言，在大笒上演奏的音高天然处于较高音区，被向下移调一个八度。相比之下，来自伽倻琴、玄琴和牙箏等弦乐器的音高在演奏中往往处于较低音区，被向上移调一个八度。诸如觱篥 和奚琴等通常在中央音域内运作的乐器则保持不变。这种方法使得在多样的器乐配置之间对音高模式进行有意义的比较成为可能。

在将井间谱 记谱转换为MusicXML格式的过程中，我们首先确立了一个基础时间单位，假定一个井间 （即记谱中的一个方格）对应于西方节奏单位中的1.5个四分音符。这一工作假设为跨曲目映射时值提供了初始参照。

然而，在韩国音乐中，井间的时间解释并不固定，而是随音乐语境而异。如上所述，一个井间既可代表由三个小拍构成的单位，也可代表由两个小拍构成的单位。在宫廷音乐中，三拍结构最为常见（见图1），但不同音乐体裁和曲目之间也存在差异。

在所分析的语料库中，两清都드리是唯一一首同时采用两拍定义和部分采用三拍定义的井间单位的乐曲。由于这两种定义对应不同的绝对时值，因此根据其底层拍节结构将两清都드리拆分为两个独立文件，以确保每个文件中的解释保持一致。

时间变异的另一个来源是长短，这是一种以固定数量井间定义的节奏框架，并有多种类型。打令长短 便是其中一种，在现代井间谱 记谱中，其时值按惯例加倍。为恢复预期的时间尺度，采用打令长短 的曲目时值乘以 ×0.5进行缩放。相比之下，与民乐 及其他几首曲目采用相同的三拍井间 单位，但在西方五线谱中记为四分音符，其时值为标准时值的三分之二。为恢复比例一致性，这些曲目的时值乘以 ×1.5进行缩放。这些差异表明，记谱时值反映的是惯例，而非绝对时间值。

因此，为了保留每首乐曲内部的节奏比例，同时实现曲目间比较，我们采用了分段时值缩放。具体而言，根据节奏结构的速度和表现特征，每首乐曲分别乘以 $\times 0.5$、 $\times 1$或 $\times 1.5$进行缩放。这种时值归一化确保了每首作品的时间结构在内部保持连贯，同时也使整个数据集能够进行一致的比较分析。

预处理流程包括以下步骤：

- **L加载 MusicXML 文件：**所有乐谱均使用 music21.converter.parse() 进行解析 .
- **筛除非音乐元素：**仅保留音符和和弦对象，排除休止符、连音线及其他元数据。
- **提取音高和时值：**从每个有效音符中提取音名（例如 C4、D4）及其对应的时值（以四分音符时值表示）。对于和弦，则将每个单独的音高视为独立音符。为便于跨乐器和跨乐曲比较，对音高和时值的数值均进行了归一化。根据各乐器的典型音区对音高进行移调（例如，大笒： -1 八度，伽倻琴、玄琴和牙箏： $+1$八度，觱篥和奚琴：不变）。对于时值，虽然通常将一个井间映射为 1.5 个四分音符，但会根据每首乐曲的节奏特征（长短）应用分段标度因子（$\times 0.5$、 $\times 1$ 或 $\times 1.5$）。
- **构建齐夫单位：**将每个（音高、时值）对视为一个独特的符号单元，并计算数据集中每个单元出现的频率。

这些齐夫单位构成了秩频分析的基本元素。通过构建提取出的单元的排名频率分布，我们评估了韩国传统音乐呈现齐夫特征的程度。

3 齐夫-曼德布罗特定律

本节将解释齐夫-曼德布罗特定律及其参数的含义，并介绍我们采用的参数值估计方法。

3.1 齐夫-曼德布罗特定律及其在双对数平面上的斜率

齐夫定律是一条经验统计原理，描述某个元素的频率与其频率排名之间的关系。George Kingsley Zipf 最初在自然语言研究中发现了这一规律[23]，此后，这一定律已在广泛的自然现象和社会现象中得到证实。从数学上来说，齐夫定律指出，频率 $f(r)$，$f \in \mathbb{Z}_{>0}$ 排名为 r 的元素$r \in \mathbb{Z}_{>0}$ 遵循以下关系

$$f(r) \propto \frac{1}{r^s}, \tag{3}$$

其中s,$s \in \mathbb{R}_{>0}$ 是缩放指数，在经验数据集中通常接近 1。在频率与排名的双对数图中，符合齐夫定律的数据会形成斜率为 $-s$ 的直线。在音乐中，齐夫定律中的“元素”或“齐夫单位”可以对应符号单元，例如音级类、节奏模式，或诸如（音高、时值）单元这样的组合表示。不过，经验数据往往偏离理想的齐夫形式，尤其是排名较低的高频元素，导致秩频图出现弯曲。为了解决这些偏差，齐夫–曼德布罗特定律引入了平移参数q,$q \in \mathbb{R}_{\geq 0}$[24]，并将公式修改为

$$f(r) \propto \frac{1}{(r+q)^s}. \tag{4}$$

当$q = 0$时，齐夫–曼德布罗特定律可简化为古典主义齐夫定律。在本研究中，我们将齐夫–曼德布罗特定律拟合于从韩国乐谱中提取的归一化齐夫单位频率分布。

对于齐夫定律，写作 $f(r) = A r^{-s}$,$A > 0$ 可得 $d\log f / d\log r = -s$,因此，双对数图上的斜率为常数 $-s$。相比之下，对于齐夫–曼德布罗特定律 $f(r) = A(r+q)^{-s}$，我们有 $d\log f / d\log r = -s r/(r+q)$，也就是说，局部斜率取决于排名 r，而齐夫定律中的斜率则具有普适性。由于局部斜率会随排名 r变化，~~该模型能够~~

捕捉到高频低秩区域中观察到的、相较于整体分布更为平缓的行为。取极限可以阐明这两种情形：

$$\lim_{r\to\infty}\frac{d\log f}{d\log r}=-s,\quad \text{for } r\gg q \tag{5}$$

$$\frac{d\log f}{d\log r}\approx -s\,\frac{r}{q}\quad \text{for } r\ll q. \tag{6}$$

因此，对于较大的排名（$r \gg q$），局部斜率趋近于 $-s$，齐夫–曼德博模型在尾部便与经典齐夫定律无法区分。对于较小的排名（$r \ll q$），斜率的绝对值降至约 sr/q，形成较平坦的“头部”，并使秩-频率图呈现明显的曲率。弯折发生在 $r \approx q$ 附近，因此，q 值越大，头部曲率越显著、延伸范围越广，同时渐近地保持齐夫斜率 $-s$。

3.2 参数估计

基于第3.1节所发现的秩频行为，我们现在量化数据分布与齐夫–曼德博模型的吻合程度。对于每个齐夫单位，我们通过拟合理论秩-频率曲线，估计归一化拟合的参数（q，s），以及原始拟合的参数（A，q，s）。

$$\begin{cases} f_{q,s}(r)=\dfrac{1}{(r+q)^s} & \text{for a normalized fit,}\\[2ex] f_{A,q,s}(r)=\dfrac{A}{(r+q)^s} & \text{for a raw fit,}\end{cases} \tag{7}$$

使其与观测数据相拟合。对于原始拟合，我们使用数据的秩-频率曲线。相比之下，我们拟合秩-归一化-频率曲线，其中频数经过归一化，使其在所有排名上的总和为 1。

为了求得各模型的最优参数值，我们使用了函数 `scipy.optimize.curve_fit`，这是 Python 中 SciPy[25] 库的一个函数，用于对用户定义的函数与数据进行非线性最小二乘拟合。该函数会返回最优参数值 (q,s) 和 (A,q,s)，使 $f_{q,s}(r)-\hat{}y$ 与 $f_{A,q,s}(r)-y$ 的残差平方和最小，其中 $\hat{}y$ 是归一化频率，y 是观测数据中齐夫单位的频率。

我们使用了信赖域反射算法，这是一种古典主义方法，可保证局部和全局收敛。[26,27] 信赖域反射算法是一种用于非线性最小二乘问题的优化方法，尤其适用于参数具有边界的情况。该算法通过迭代方式围绕当前估计值定义一个“信赖域”——即模型在其中被认为能够良好近似目标函数的区域——然后在该区域内寻找一个可减小残差的步长。每一步之后，算法会根据模型对改进效果的预测情况扩大或缩小该区域。在 scipy.optimize.curve_fit 中，当为参数提供边界（通过 bounds 参数）时，会自动使用信赖域反射算法。如果未指定边界，默认算法为莱文贝格–马夸尔特算法，该算法无法处理参数边界。对于我们的情况，边界设为 $0 \le q \le 1000, 0 \le s \le 20$ 以防止溢出。

4 数据分析

4.1 统计摘要

我们首先汇总数据集中出现频率最高的单元（表1），以提供基本的描述性统计。本节中，每个音符的时值均以单元为单位进行测量，其中数值 1。5 对应一个井间。单元总数为 65，817，因此我们以百分比报告相对频数，并四舍五入至小数点后一位。表1 汇总了语料库中频数最高的音高、时值以及（音高、时值）单元。

音高。排名最高的音高集中在中央音域附近的一个八度内：排名前七的音高类型均处于这一音域，表明音高分布集中于同一八度内的一小组音。其中特别常见的是 E♭4 和 B♭4，通常作为持续的主音演奏；而 F4 则常作为过渡音，连接相邻的主音。

时值。时值单位主要为 1.5、0.5 和 1.0。按照我们的约定，1.5 对应一个井间。由于宫廷音乐中一个井间最常采用三拍结构（见图1），数值 0.5 和 1.0 自然对应一个井间的一个拍和两个拍细分，这有助于解释它们为何频数较高。

（音高、时值）。这种结构也延续到了联合单元中：出现频率最高的音高—时值对包括（E♭4，1.5）和（B♭4，1.5），以及（F4，0.5）。尤其是（F4，0.5）的突出表现，与过渡音时值较短的倾向一致；而主音则更常以较长、持续一个完整井间的时值出现。

表1. 排名最高的音高 音高、时值和（音高、时值）单元。排名在各类别内分别计算 y.

Rank	音高			时值			（音高、时值）		
	Unit	计数	比率(%)	Unit	计数	比率(%)	Unit	计数	比率(%)
1	E♭4	10177	15.5	1.5	14297	21.7	E♭4, 1.5	2645	4.0
2	B♭4	9884	15.0	0.5	14272	21.7	F4, 0.5	2404	3.7
3	F4	9440	14.3	1.0	10351	15.7	B♭4, 1.5	2390	3.6
4	A♭4	8848	13.4	0.25	7706	11.7	A♭4, 1.5	2025	3.1
5	B♭3	5588	8.5	3.0	5985	9.1	B♭4, 1.0	1966	3.0
6	C4	3953	6.0	0.75	3160	4.8	B♭4, 0.5	1965	3.0
7	C5	3024	4.6	1/12	2319	3.5	F4, 1.5	1690	2.6
8	E♭5	2973	4.5	2.5	2231	3.4	E♭4, 0.5	1682	2.6
9	A♭3	2159	3.3	1/4	1775	2.7	A♭4, 1.0	1546	2.3
10	G4	2088	3.2	1/3	1391	2.1	A♭4, 0.5	1533	2.3

表2. 联合层面与乐器层面的齐夫–曼德布罗特拟合结果。

面板A：联合层面拟合。

Unit	Fit	A	q	s	R^2	N	$bar_{\min}$	$bar_{\max}$
(p,d)	raw	66856.82	41.46	2.64	0.9980	555	18	34
(p,d)	norm	1.00	43.88	2.76	0.9979	555	18	33
(p)	raw	69222.83	101.61	20.00	0.9620	30	5	6
(p)	norm	1.00	96.90	20.00	0.9600	30	5	6
(d)	—	—	—	—	—	—	—	—

仅时值拟合在相同的拟合流程下未能收敛。

B面板：乐器层级拟合。

乐器	Fit	A	q	s	R^2	曲目	N	L
玄琴	raw	7433.615	31.774	2.995	0.994	30	184	7330
玄琴	norm	1.000	33.753	3.147	0.994	30	184	7330
Piri	raw	13 894.548	77.050	4.451	0.996	46	289	14 012
Piri	norm	1.000	73.205	4.257	0.996	46	289	14 012
伽倻琴	raw	7807.929	42.090	3.796	0.991	30	158	7708
伽倻琴	norm	1.000	45.330	4.042	0.991	30	158	7708
Song	raw	912.678	76.275	3.058	0.943	3	166	904
Song	norm	1.000	80.221	3.194	0.943	3	166	904
大笒	raw	20 522.660	61.867	2.855	0.994	46	387	19 985
大笒	norm	1.000	69.229	3.129	0.994	46	387	19 985
奚琴	raw	13 011.304	56.688	3.891	0.997	46	236	12 972
奚琴	norm	1.000	57.586	3.945	0.997	46	236	12 972
牙箏	raw	2940.586	2.343	1.115	0.964	15	121	2731
牙箏	norm	1.000	2.786	1.217	0.962	15	121	2731

4.2 实验设置与概述

我们主要将（音高、时值）音高—时值对作为齐夫单位来分析齐夫–曼德布罗特拟合，并另外报告仅音高和仅时值拟合，作为诊断性比较。我们通过固定$A = 1$来考察归一化拟合，并通过将A视为方程7中的自由参数，考察原始拟合。

• **全局合并数据**。在所有乐曲中，我们考虑了三种齐夫单位：按顺序排列的（音高、时值）对、单独的音高，以及单独的时值。

• **乐器级合并数据**。对于每个乐器类别（玄琴、觱篥、伽倻琴、大笒、奚琴、牙箏和人声），我们构建了针对各乐器的合并数据集。在这一层级，我们只关注（音高、时值）对，因为单一乐器的样本量不足以对仅音高或仅时值单元进行稳定的 Zipf 拟合。

• **乐曲级数据**。 对于每首乐曲（完整曲目列表记录于表 3），我们再次使用 (音高, 时值) 音高—时值对作为齐夫单位进行齐夫–曼德布罗特拟合。

为了量化拟合的齐夫–曼德布罗特曲线与观测到的秩-频率分布的吻合程度，我们采用了R^2 分数。令$y_r = \log f(r)$ 表示排名r处观测到的对数频率，$\hat{y}_r = \log \hat{f}(r)$ 表示对应的拟合值，其中$r = 1,\ldots,N$。我们计算

$$R^2 = 1 - \frac{\sum_{r=1}^{N}(y_r - \hat{y}_r)^2}{\sum_{r=1}^{N}(y_r - \bar{y})^2}, \qquad \bar{y} = \frac{1}{N}\sum_{r=1}^{N} y_r,$$

这与皮尔逊相关系数密切相关，可解释为拟合曲线对 $\{y_r\}$ 方差的解释比例[28]。我们主要将R^2 用作描述性相似度指数，以比较不同单元和范围内的众多拟合结果，这与此前音乐和语言领域的齐夫定律/齐夫–曼德布罗特研究一致[29,30]。虽然没有普遍适用的标准来界定何为“良好”拟合，但高于 0.7 的数值通常被视为表明拟合良好；0.8 和 0.9 则表示拟合程度更高[31–33]。主要发现如下：

- **全局并集**。（音高、时值）单元与齐夫–曼德布罗特分布高度吻合，其R^2得分为 0.9980（原始拟合）和 0.9979（归一化拟合）（图3（a）、（b），表2（面板A））；而仅音高拟合的参数达到边界，仅时值拟合则未能收敛（图3（c）、（d））。

- **乐器级并集**。拟合仍然表现良好，原始拟合和归一化拟合的平均R^2分数为 0。982，但尾部行为因乐器而异（图4，表2（B面板））。

- **乐曲级**。大多数曲目都可以进行稳定拟合，归一化拟合的平均R^2分数为 0。978（表3）。

详细结果将在第4.3 - 4.5节讨论。

表2–3的阅读方法。这些表格采用统一的记谱方式。“原始”和“归一化”分别表示原始齐夫–曼德布罗特拟合和归一化拟合，其中$A = 1$。拟合参数为A、q、s和R^2是第R^2节所述的分数4.2。这里N 表示不同齐夫单位的数量，即最大排名；L 表示相应数据集中具有重数的齐夫单位总数。

4.3 全局并集级拟合

对于全局合并数据，当（音高，时值）音高—时值对被用作齐夫单位时，齐夫–曼德博模型对秩-频率分布的拟合效果极佳。详细的参数估计见表2 （面板A），相应的双对数图见图3（a）–（b）。

图像均采用双对数尺度绘制。每幅图中，绿色阴影带标示了拟合的齐夫–曼德布罗特曲线斜率介于 -1.2 和 -0.8 之间的排名范围。相应的下限和上限秩索引以$_{\min}$ 和$_{\max}$ 的形式列于表中。存在这样一个斜率接近 -1 的近线性区域，表明拟合模型成功捕捉了广义齐夫定律。

当仅将音高或时值用作齐夫单位时，模型要么无法收敛（时值），要么虽得到形式上较高的R^2，但参数却被推至优化边界（音高），这表明拟合没有实际意义。当将（音高，时值）音高—时值对用作齐夫单位时，归一化拟合和原始拟合的R^2 均达到约0.998。此处，归一化拟合将$A = 1$ 固定，并用频数除以节点总数；原始拟合则将A 作为自由参数。

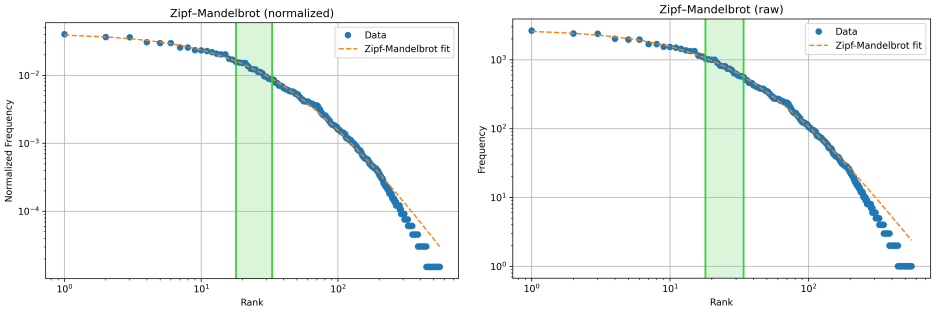
当我们选择时值作为齐夫单位时，模型无法收敛，也无法找到最优解。另一方面，如果选择音高作为齐夫单位，音高分布的头部较为平坦，难以遵循齐夫-曼德布罗特定律，导致s 达到其上界。尽管R^2 得分很高，我们推断这并非有意义的拟合。我们将在第5.1节对此进行详细讨论。

当我们将（音高、时值）对视为齐夫单位时，发现模型总体上很好地拟合了数据，但未能成功拟合尾部；实际数据呈现出较弱的尾部。在按单个乐器拟合数据时，我们发现了一个有趣的现象。

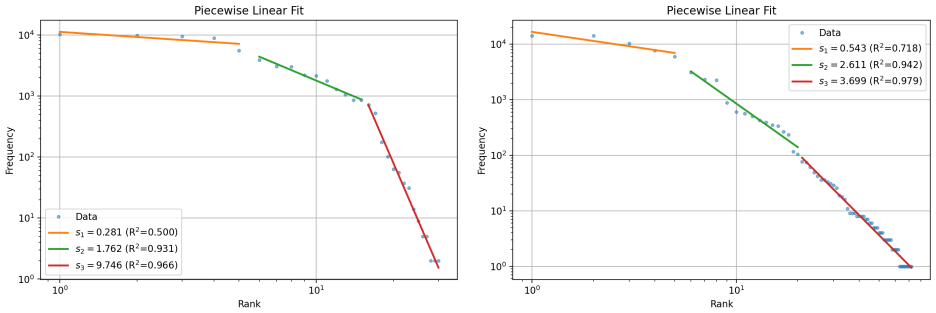
4.4 乐器层级拟合

受图3(a)–(b)中联合层级结果的启发，我们接下来按乐器对数据进行分层。在联合拟合中，中排名区域的吻合度很高，但远尾部出现偏差：观测频数低于拟合的齐夫–曼德布罗特曲线，即对于较大的排名$f(r) < \hat{f}(r)$ 时r。这表明联合分布可能是依赖乐器的秩频机制的混合，而非单一均质机制。

对于每种乐器类别，我们构建乐器层级联合数据集，并使用（音高、时值）音高—时值对作为齐夫单位来拟合齐夫–曼德博模型。图4 并列展示了归一化拟合与原始拟合，表2（B面板）总结了拟合参数和样本量。不同乐器的曲目数量和观测总数差异显著：例如，大笒贡献了最多的数据量（$L = 19$，共985，分布于46首曲目），其次是鼜篥（$L = 14$，共012，46首曲目）和奚琴



(a) 并集层级：归一化拟合　　**(b)** 并集层级：原始拟合



(c) 仅音高：分段线性拟合（3段）　　**(d)** 仅时值：分段线性拟合（3段）

图3. 针对（音高、时值）单元的并集层级齐夫–曼德博拟合，以及边际仅音高与仅时值秩-频率曲线的分段线性拟合（双对数尺度）。

（$L = 12$，共972，46首曲目），而人声仅由3首曲目代表（$L = 904$），因为每首乐曲都由不同的乐器家族演奏。

总体而言，各乐器的拟合仍然很强，R^2 通常处于0.962 – 0.997（表2，面板B）的范围内，表明齐夫–曼德布罗特曲线捕捉了主要的秩频趋势。然而，尾部行为明显因乐器而异。大多数乐器在整个阴影中段斜率区域直至尾部，数据与拟合之间均表现出密切一致。相比之下，大笒与牙箏表现出尾部缺失：在高排名处，拟合曲线高估了经验频率（$f(r) < \hat{f}(r)$），这意味着稀有（音高、时值）单元的出现次数少于拟合的齐夫–曼德博模型所预测的次数。同样的定性模式在全局并集拟合中可见，表明乐器层级异质性是并集层级尾部失配的一个合理来源。

对于牙箏，整体吻合度相对较弱（$R^2 \approx 0.962$），经验曲线的中排名区域明显更为平坦，许多（音高、时值）类型的频数非常相近。从音乐角度来看，这对应于一种将使用更均匀地分布在各单元上的分布，因此单个旋律单元不那么突出，乐句往往更多地发挥背景织体的作用。

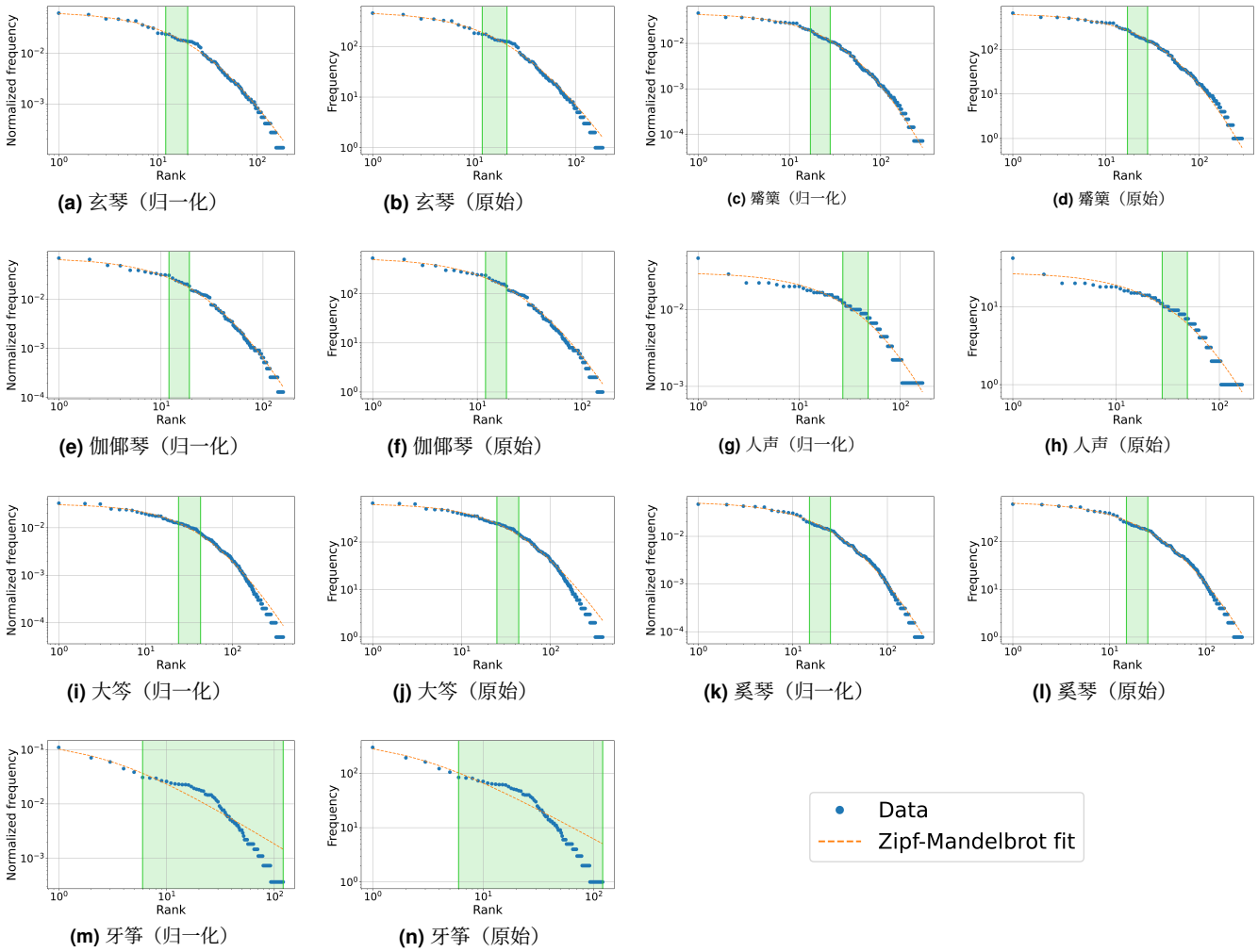
对于人声，较低的R^2 (0.943)也应谨慎解读，因为它基于数量较少的曲目和小得多的样本量。尽管如此，头部排名靠前的少数几项之后频数便急剧下降，这与人声 声部往往高度依赖少量骨干音的倾向相符。

综合来看，乐器层级拟合支持这样一种解释：全局并集曲线汇总了多种乐器特有的模式。特别是，不同乐器偏离齐夫–曼德博尾部的方式各不相同（大笒的偏离程度较轻但涉及较大音量，牙箏则表现出更明显的结构性偏离）。

4.5 乐曲层级拟合

接下来，我们考察在联合层级观察到的齐夫–曼德布罗特规律是否也存在于各首乐曲中。表3 报告了每首乐曲的归一化齐夫–曼德博拟合，其中以(音高，时值)对作为齐夫单位。

总体而言，乐曲层级吻合度始终较高。在语料库的 46 首曲目中，R^2 值介于 0.930 和 0.995 之间（中位数为 0.985），且所有曲目均满足 $R^2 \geq 0.9$。此外，大多数曲目都超过了常用的“高度吻合”阈值：89% 的曲目满足 $R^2 \geq 0.95$，83% 的曲目满足 $R^2 \geq 0.97$。这些结果表明，齐夫–曼德布罗特结构并非仅存在于全局联合语料库中，而是已经体现在单首作品层面。



(a) 玄琴（归一化） (b) 玄琴（原始） (c) 箫篥（归一化） (d) 箫篥（原始）

(e) 伽倻琴（归一化） (f) 伽倻琴（原始） (g) 人声（归一化） (h) 人声（原始）

(i) 大笒（归一化） (j) 大笒（原始） (k) 奚琴（归一化） (l) 奚琴（原始）

(m) 牙箏（归一化） (n) 牙箏（原始）

图4. 乐器层面的 Zipf –Mandelbrot 拟合，针对（音高、时值）单位（双对数尺度）。对于每种乐器，归一化 fit（左）与原始拟合（右）并列显示 n。

与此同时，各首乐曲的拟合参数（q,s）各不相同，反映出秩-频率曲线形态因乐曲而异（例如，头部向中段过渡的速度，以及尾部的厚重程度）。不同乐曲中不同（音高、时值）单元的数量也有很大差异（N 介于 47 至 213 之间），单元总数亦然（L 介于 380 至 3,229 之间）。我们观察到，最低的 R^2 值往往出现在较短的乐曲中。

许多曲目都能取得良好的拟合，这种一致性表明，排名–频率组织是该曲目库的一项稳健特征，而非联合本身的属性。一种可能的解释是，该语料库具有共同的作曲特征和风格惯例，并在各曲目中反复体现。

5 关于头部行为与联合分布的说明

5.1 为什么平坦头部会导致拟合失败?

考虑齐夫–曼德布罗特模型 $f(r) = A(r+q)^{-s}$，以及它在双对数图上的局部斜率 :

$$\frac{d\log f}{d\log r} = -s\,\frac{r}{r+q}. \tag{8}$$

假设经验头部在范围 $1 \le r \le r_h$ 内（近乎）平坦，这可以表示为 $\left|{}^{d}\log {}_{fd}\log_r\right| = s\,r\,r+q \le \varepsilon.$ 于是，

$$q \;\ge\; \left(\frac{s}{\varepsilon}-1\right) r, \text{ for } 1 \le r \le r_h.$$

因此，要让头部在较宽的观测范围内保持足够平坦（r_h较大），平移参数 q 必须较大。从几何角度看，这会将弯曲点（大约出现在 $r \approx q$处）向右移动。根据公式 8，当 $r \ll q$时，斜率趋近于 0；而当 $r \gg q$时，斜率趋近于 $-s$。因此，较大的 q 可确保观测区域（即有观测值的数据部分，通常为 $1 \le r \le N$）仍处于 $r \ll q$ 范围内。

由于真实数据集具有有限范围 $1 \le r \le N$，将q 取得非常大，可使$r r+q$ 在整个观测区域内保持很小，从而消除弯曲点。这使得难以可靠地识别向尾部斜率 $-s$ 的过渡。

然而，如果弯曲点存在于观测范围之内（$1 \le r \le N$），拟合参数s 往往会同时增大以进行补偿，因为斜率$\frac{d\log f}{d\log r} = -s\, r r+q$ 仍必须拟合数据的曲率。在实践中，表2（图A）说明了当音高被用作齐夫单位时的这一行为：拟合值q 为96.90，而$N = 30$，这意味着整个观测范围都位于$r \ll q$ 情形内。因此，估计的s 达到其上界（$s = 20$）。

5.2 两个独立齐普夫单元的联合分布

我们观察到，（音高，时值）单元的秩-频率分布与从我们的数据集所导出的概率分布中独立抽样音高和时值而得到的音高—时值对几乎相同，报告的R^2得分为0.993。虽然（音高，时值）单元的秩-频率分布遵循齐夫-曼德布罗特定律，但我们发现音高与时值单元各自的秩-频率分布均遵循具有3条线段的分段齐夫定律。基于这一观察，本节验证两个遵循齐夫定律的独立分布的联合分布是否遵循齐普夫-曼德尔布罗特模型。为方便起见，我们将分段齐普夫模型简化为齐普夫模型。设

$$f_{X_1}(x_1) = x_1^{-t_1}, \qquad f_{X_2}(x_2) = x_2^{-t_2}, \qquad t_1, t_2 > 0,$$

$$f_{X_i} : X_i \to \mathbb{R}_{\ge 0}, \qquad X_i = [1, \infty), \qquad i = 1, 2,$$

并进一步假设X_1 和X_2 相互独立。其联合密度函数为

$$f(x_1, x_2) = f_{X_1, X_2}(x_1, x_2) = x_1^{-t_1} x_2^{-t_2}, \qquad (x_1, x_2) \in [1, \infty)^2.$$

对于阈值$f_0 \in (0, 1]$, ，定义上水平集f_0 为

$$\Omega(f_0) := \big\{(x_1, x_2) \in [1, \infty)^2 : \ f(x_1, x_2) \ge f_0\big\},$$

其面积A： $(0, 1] \to \mathbb{R}_{\ge 0}$ 为

$$A(f_0) := \mathrm{Area}\big(\Omega(f_0)\big) = \iint_{\Omega(f_0)} dx dy.$$

由于f 在每个坐标方向上都严格递减，A 在$(0, 1]$ 上也严格递减，因此可逆。

现在我们要将函数A 与联合分布的秩—频率分布联系起来。对于$N_1, N_2 \in \mathbb{N}$，定义 $[1, \infty)^2$ 中的格为

$$\mathscr{L}(N_1, N_2) := \left\{ \left(\frac{m}{N_1}, \frac{n}{N_2} \right) \in [1, \infty)^2 : \ m, n \in \mathbb{N} \right\}.$$

由于f 在每个坐标方向上都严格递减，我们可以找到以下集合中第r大的元素（该元素在$r \times r$ 格中必然存在）：

$$\Big\{ f(x_1, x_2) \ \Big| \ (x_1, x_2) \in \mathscr{L}(N_1, N_2) \Big\}$$

定义 $f_{*(N_1, N_2)}$： ${}^{1}_{N_1 N_2}\mathbb{N} \to \mathbb{R}_{>0}$ 为

$$\frac{r}{N_1 N_2} \longmapsto \text{the } r\text{-th largest value of } \Big\{ f(x_1, x_2) \ \Big| \ (x_1, x_2) \in \mathscr{L}(N_1, N_2) \Big\},$$

其中集合 ${}^{1}_{N}\mathbb{N}$： $= \{ \frac{n}{N} \mid n \in \mathbb{N} \}$.

备注 5.1。若 $N_1 = N_2 = 1$，则函数 $f_{*(1,1)}$ 恰好与联合分布的秩-频率图相同。

对于 $f_0 \in (0, 1]$, ，定义归一化离散计数泛函

$$A_{N_1, N_2}(f_0) := \frac{1}{N_1 N_2} \#\big(\mathscr{L}(N_1, N_2) \cap \Omega(f_0)\big).$$

最后，我们说明 $f_{*(N_1, N_2)}$ 与 A^{-1} 之间的联系。详细证明见附录 B。

命题 5.2（排序收敛至反面积曲线）。 固定 $r > 0$ 和 $\varepsilon_1, \varepsilon_2 > 0$。那么存在 $M_1, M_2 \in \mathbb{N}$，使得对于每个 $N_1 \geq M_1$ 以及 $N_2 \geq M_2$，都存在一个整数 $c \in \mathbb{N}$，使得

$$\left| r - \frac{c}{N_1 N_2} \right| < \varepsilon_2 \qquad \textit{and} \qquad \left| A^{-1}(r) - f_{*(N_1,N_2)}\left(\frac{c}{N_1 N_2}\right) \right| < \varepsilon_1.$$

特别地，曲线 $\{(r, A^{-1}(r)) \mid r \geq 0\}$ 是将格点样本按 f 降序排序后得到的秩频图的连续类比。

在命题 5.2中，我们已经证明，联合分布 $f(x_1, x_2) = x_1^{-t_1} x_2^{-t_2}$的秩-频率曲线恰好是函数 $f_*(r) = A^{-1}(r)$。要检验这条曲线是否能用齐夫–曼德布罗特定律很好地近似，必须显式计算 $A(f_0)$，接下来我们将进行计算。同样，详细证明见附录 B。

命题 5.3（排序函数的闭式表达式）。 固定 $t_1, t_2 > 0$。对于 $f_0 \in (0, 1]$, 定义上水平集

$$\Omega(f_0) := \left\{ (x_1, x_2) \in [1, \infty)^2 : x_1^{-t_1} x_2^{-t_2} \geq f_0 \right\},$$

并令 $A(f_0) :=$ 面积 $\Omega(f_0)$。则：

1. 如果 $t_1 \neq t_2$,

$$A(f_0) = 1 + \frac{t_1 f_0^{-1/t_1} - t_2 f_0^{-1/t_2}}{t_2 - t_1}.$$

2. 若 $t_1 = t_2 = t$,

$$A(f_0) = f_0^{-1/t}\left(\tfrac{1}{t} \ln \tfrac{1}{f_0} - 1\right) + 1.$$

此外，$A : (0, 1] \to [1, \infty)$ 严格递减，因此是双射；连续秩频曲线是其反函数

$$f_*(r) = A^{-1}(r) \qquad (r \geq 1).$$

得到$A(f_0)$的显式公式后，对于任意 (t_1, t_2)，都可以计算秩-频率曲线$f_*(r) = A^{-1}(r)$。现在，我们通过对 $\{(r, f_*(r))\}$ 拟合齐夫-曼德布洛特模型，并报告由此得到的R^2 分数，来检验理论计算的曲线是否能被齐夫–曼德布罗特定律很好地近似。

验证。给定 (t_1, t_2)，我们使用 scipy.optimize.curve_fit 对 $\{(r, f_*(r))\}$ 拟合齐夫-曼德布洛特模型，从而得到 $(A, q\ \alpha)$。验证针对 (t_1, t_2) 进行，取值范围为 0.5 到 4，步长为 0.5。我们还计算了每次拟合的 R^2 分数。结果记录在表 4 中。

上述结果具有若干重要的理论意义。首先，它可能解释重尾行为在配对或联合构造下的稳定性。齐夫定律刻画的是边缘秩-频率分布，而许多现实世界系统涉及两个或更多齐夫单位之间的相互作用。证明它们的联合分布遵循齐夫-曼德布罗特定律，表明齐夫-曼德布罗特定律可能自然地作为齐夫型标度的高阶扩展形成。该结果还有助于解释齐夫定律与齐夫-曼德布罗特定律之间的关系。后者是齐夫定律的一种经验性扩展，增加了一个平移参数。我们的结果可能揭示齐夫-曼德布罗特定律出现的一种有原理依据的生成方式。据作者所知，尚无先前研究从理论上考察独立齐夫型数据的联合分布如何促成齐夫-曼德布罗特定律的出现。这一主题值得进一步研究。

6 结论

齐夫定律及其扩展形式齐夫–曼德布罗特定律，常见于各种自然数据中，包括自然语言和音乐。尤其是音乐中的齐夫现象，表明音乐的作曲本质具有某种语义结构。这意味着，人类所感知到的音乐审美特征可能源于数据中潜在的标度模式。事实上，各种研究表明，齐夫–曼德布罗特定律在不同时期和不同流派的音乐数据中广泛显现，尤其是在西方音乐中。

本文提出了两项重要发现。首先，我们分析了以传统记谱系统井间谱记写的韩国传统音乐，即被称为 正乐 的宫廷音乐。我们发现，韩国传统音乐显著遵循齐夫–曼德布罗特定律，将音高、时值及其配对组合用作齐夫单位。对于每个单位，都观察到了清晰的

齐夫–曼德布罗特标度行为。此外，这种标度律似乎比在西方音乐中观察到的更为显著。一种可能的解释是，韩国传统音乐数百年来不仅受到个体作曲家的塑造，也在社会中由集体共同塑造。因此，其作曲结构可能以一种广泛为社群成员所接受的方式演化，从而呈现出更加明显的符合定律的模式。我们的研究结果进一步证明，音乐数据在不同文化和历史时期中都呈现出齐夫特性。

本文还表明，两个独立齐夫型数据集的联合分布会形成遵循齐夫–曼德布罗特定律的分布。这或许可以解释为何齐夫–曼德布罗特定律会作为齐夫定律的扩展而出现，因为大多数自然数据都源自两个或更多齐夫单位的联合分布。事实上，这一结果也适用于我们的韩国音乐数据。具体而言，音高和时值各自都符合分段齐夫定律，而将它们组合成成对单元后，便会得到齐夫–曼德博分布。据作者所知，尚无先前研究从理论上探讨独立齐夫型数据的联合分布如何导致齐夫–曼德布罗特定律的出现。

本研究聚焦于正乐数据。正乐是韩国传统音乐的主要类别之一，属于宫廷音乐。另一重要类别是韩国传统民间音乐，其风格通常更为灵活，且并非始终以正式乐谱记谱。我们计划进一步探究这一类别中的齐夫–曼德布罗特定律。我们推测，齐夫–曼德布罗特定律之所以表现得如此显著，可能是因为数百年来，特有的变异逐渐被平滑消解，最终形成了由集体文化共识塑造的音乐形式。我们将在未来的工作中进一步探讨这一假设。

关于联合分布的讨论，我们的理论分析仅限于两个独立齐夫型数据集的联合分布，每个数据集都由单一齐夫斜率表征。我们将把这项研究扩展到分段齐夫数据集的组合。此外，目前的工作仅考虑两个齐夫数据集的联合分布。未来的研究将考察多个数据集，并探讨齐夫–曼德布罗特定律更一般的框架。

参考文献

1. Zipf, G. K. 《语言的心理生物学：动态语文学导论》（Houghton Mifflin，1935）。2. Zipf, G. K. 《人类行为与最省力原则》（Addison-Wesley，1949）。3. Mandelbrot, B. 语言统计结构的信息理论。《通信理论》486–502（1953）。4. Ferrer-i Cancho, R. 与 Solé, R. V. 最省力原则与人类语言标度的起源。《美国国家科学院院刊》100，788–791，DOI: 10.1073/pnas.0335980100（2003）。5. Hackmann, A. 美国城市齐夫定律的演变。《物理学A辑：统计力学及其应用》563，125397（2020）。6. Adamic, L. A. 与 Huberman, B. A. 齐夫定律与互联网。《语言计量学》3，118–120（2002）。7. Frank, S. A. 对数级数与齐夫定律。《理论生物学杂志》461，1–9（2019）。8. Mandelbrot, B. 《信息论与心理语言学：词语与大众文化理论》（Mouton，1965）。9. Herdan, G. 《形符—类符数学》（Mouton，1960）。10. Voss, R. F. 与 Clarke, J. 音乐中的1/f 噪声：由1/f 噪声生成的音乐。《美国声学学会杂志》63，258–263（1978）。11. Manaris, B. 等. 齐夫定律、音乐分类与美学。《计算机音乐杂志》29，55–69（2005）。12. Perotti, J. I. 与 Billoni, O. V. 音乐中齐夫定律的涌现。《物理学A辑：统计力学及其应用》545，123546（2020）。13. Zanette, D. H. 齐夫定律与音乐语境的形成。《物理学A辑：统计力学及其应用》337，307–314（2004）。14. Simon, H. A. 论一类偏态分布函数。《生物统计学》42，425–440（1955）。15. Serra-Peralta, M. 等. 古典音乐史中的希普斯定律与词汇丰富度。《欧洲物理学杂志数据科学》10，12（2021）。16. Haro, M., Serrà, J., Herrera, P. 与 Corral, Á. 语音、音乐与环境声音信号短时音色编码中的齐夫定律。《公共科学图书馆·综合》7，e33993，DOI: 10.1371/journal.pone.0033993（2012）。17. Kim, M. o. 与 Jung, J.-h. 国乐的科学理解与人工智能国乐II：用于井间谱数据构建的正乐曲目集。KIAS跨学科研究丛书（Yesol，首尔，2023）。18. Corral, Á., Ferrer-i Cancho, R. 与 Boleda, G. 词频的齐夫定律：长文本中的词形与词元。《公共科学图书馆·综合》10，e0129031，DOI: 10.1371/journal.pone.0129031（2015）。19. Good, M. 用于记谱与分析的MusicXML。《虚拟乐谱：表示、检索与修复》12，160（2001）。20. Cuthbert, M. S. 与 Ariza, C. music21：计算机辅助音乐学与符号音乐数据工具包。《国际音乐信息检索学会会议论文集》（2010）。

21. Hwang, C. Y. 《灵山会相》研究（首尔大学出版社，1999）。22. Chun, I. P. 韩国乐器音乐创作指南（Heanday Music，2013）。23. Zipf, G. K. 人类行为与最省力原则：人类生态学导论（Ravenio books，2016）。24. Cannon, J. 自然的分形几何。Benoit B. Mandelbrot著。《美国数学月刊》91，594–598（1984）。25. Virtanen, P. 等. SciPy 1.0：Python科学计算的基础算法。《自然·方法》17，261–272（2020）。26. Branch, M. A., Coleman, T. F. 与 Li, Y. 面向大规模有界约束最小化问题的子空间、内点与共轭梯度法。《SIAM科学计算杂志》21，1–23（1999）。27. Byrd, R. H., Schnabel, R. B. 与 Shultz, G. A. 通过在二维子空间上最小化近似求解信赖域问题。《数学规划》40，247–263（1988）。28. Pearson, K. VII. 对进化论的数学贡献。——III. 回归、遗传与泛交。《伦敦皇家学会哲学汇刊》A辑，收录数学或物理学性质的论文，253–318（1896）。29. Liu, L., Wei, J., Zhang, H., Xin, J. 与 Huang, J. 从巴赫到肖邦的古典音乐中音高波动的统计物理学视角：标度律的证据。《公共科学图书馆·综合》8，e58710（2013）。30. Marques, V. M. 与 Reis, C. 识别音乐风格——一种基于齐夫-曼德布罗特定律的方法。载于《计算智能与决策》：趋势与应用，439–448（施普林格，2012）。31. Manaris, B., Purewal, T. 与 McCormick, C. 用计算机识别和分类优美音乐的进展——MIDI编码音乐与齐夫-曼德布罗特定律。载于IEEE东南会议2002论文集（分类号02CH37283），52–57（IEEE，2002）。32. Manaris, B., Vaughan, D., Wagner, C., Romero, J. 与 Davis, R. B. 进化音乐与齐夫-曼德布罗特定律：为悦耳音乐开发适应度函数。载于《进化计算应用研讨会》，522–534（施普林格，2003）。33. Linders, G. M. 与 Louwerse, M. M. 再论齐夫定律：口语对话、语言单位、参数与最省力原则。《心理通报与评论》30，77–101（2023）。34. Kneser, M. 关于闵可夫斯基面积测度的若干评注。《数学档案》6，382–390（1955）。

致谢

本研究由韩国国家研究基金会（NRF）通过项目编号 2021R1A2C3009648 资助，并部分获得 KIAS 跨学科研究计划的支持。

附录

按音域划分的律名 记谱

Yulmyeong	pitch	Yulmyeong	pitch	Yulmyeong	pitch
僙	E♭3	黃	E♭4	潢	E♭5
仗	F3	太	F4	汰	F5
伸	A♭3	仲	A♭4	沖	A♭5
侻	B♭3	林	B♭4	淋	B♭5
倘	C4	南	C5	滴	C6

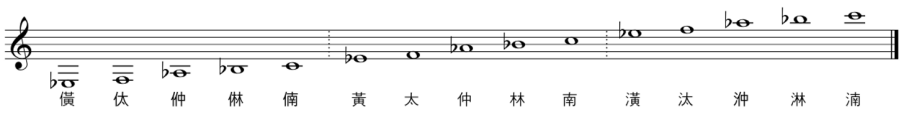


图 5. 按八度应用音名

按音域对律名 的记谱：现代井间谱 使用宽广的音符音域。因此，在律名 上添加汉字以区分音符音域。这在各种器乐声部中普遍使用。这些不同的音符音域均可转换为西式音符。

B 命题 5.2 和5.3 的数学证明

命题 B.1（排序收敛于反面积曲线）。 固定 $r > 0$ 与 ε_1、$\varepsilon_2 > 0$。则存在 $M_1, M_2 \in \mathbb{N}$，使得对每一个 $N_1 \geq M_1$ 与 $N_2 \geq M_2$，存在整数 $c \in \mathbb{N}$，使得

$$\left| r - \frac{c}{N_1 N_2} \right| < \varepsilon_2 \qquad \text{and} \qquad \left| A^{-1}(r) - f_{*(N_1,N_2)}\left(\frac{c}{N_1 N_2}\right) \right| < \varepsilon_1.$$

特别地，曲线 $\{(r, A^{-1}(r)) \mid r \geq 0\}$ 是将格点样本按 f 降序排序后得到的秩频图的连续类比。

证明。 **步骤 1**。 固定 $f_0 \in (0,1]$ 并记 $\Omega = \Omega(f_0)$。令

$$U := f_0^{-1/t_1}, \qquad V := f_0^{-1/t_2},$$

所以 $\Omega \subset [1,U] \times [1,V]$。

对于整数 m，$n \in \mathbb{N}$，定义网格间距为 h_1 的网格矩形：$= {}^{1}\!/_{N1}$，$h_2 := {}^{1}\!/_{N2}$ 如下

$$Q_{m,n} := \left[\frac{m}{N_1}, \frac{m+1}{N_1}\right] \times \left[\frac{n}{N_2}, \frac{n+1}{N_2}\right].$$

考虑外部网格近似

$$\Omega^{+}_{N_1,N_2} := \bigcup_{\left(\frac{m}{N_1}, \frac{n}{N_2}\right) \in \Omega} Q_{m,n}.$$

那么，联合中的矩形 $Q_{m,n}$ 内部互不相交，每个矩形的面积为 ${}^{1}\!/_{N1N2}$，且上述指标集恰好为 $\mathscr{L}(N_1, N_2) \cap \Omega$。因此

$$\mathrm{Area}(\Omega^{+}_{N_1,N_2}) = \frac{1}{N_1 N_2} \#\left(\mathscr{L}(N_1,N_2) \cap \Omega\right) = A_{N_1,N_2}(f_0).$$

此外 $\Omega \subset \Omega^{+}_{N_1}$，对于给定的 N2：$(x_1, x_2) \in \Omega$，令 $m = \lfloor N_1 x_1 \rfloor$，$n = \lfloor N_2 x_2 \rfloor$。则 $(m/N_1, n/N_2) \leq (x_1, x_2)$ 按坐标逐项成立。由于 f 在每个坐标上均递减，$f(m/N_1, n/N_2) \geq f(x_1, x_2) \geq f_0$,，因此 $(m/N_1, n/N_2) \in \Omega$，从而 $(x_1, x_2) \in Q_{m,n} \subset \Omega^{+}_{N_1}$,N2。

类似地，定义内部网格近似

$$\Omega^{-}_{N_1,N_2} := \bigcup_{\left(\frac{m+1}{N_1}, \frac{n+1}{N_2}\right) \in \Omega} Q_{m,n}.$$

如果 $Q_{m,n}$ 的右上角位于 Ω，那么 $Q_{m,n}$ 中的每个点的坐标都更小，因此 f 值更大，所以 $Q_{m,n} \subset \Omega$。因此 $\Omega^{-}_{N_1,N_2} \subset \Omega \subset \Omega^{+}_{N_1,N_2}$。

现在，$\partial\Omega$ 是分段 C^1 曲线（图像为 $x_1 = f^{-1/}t10\, x^{-t_2/}t12$，并在 $x_1 = 1$ 和 $x_2 = 1$上包含线段），因此具有有限长度；记为 $P(f_0)$。任何在 $\Omega^{+}_{N_1}$,$N_2 \setminus \Omega^{-}_{N_1}$,N2 中计数的矩形都必须与 $\partial\Omega$相交，且包含在 $\partial\Omega$的 $\sqrt{2}h$-邻域内，其中 $h := \max\{h_1, h_2\}$。根据克内瑟定理，管状邻域[34], 的面积为 O(h)；具体而言，

$$0 \leq \mathrm{Area}(\Omega^{+}_{N_1,N_2}) - \mathrm{Area}(\Omega^{-}_{N_1,N_2}) \leq 2P(f_0)\,h + O(h^2) \xrightarrow[N_1,N_2 \to \infty]{} 0.$$

由于 $\Omega^{-}_{N_1,N_2} \subset \Omega \subset \Omega^{+}_{N_1,N_2}$，我们得到

$$\mathrm{Area}(\Omega^{+}_{N_1,N_2}) \to \mathrm{Area}(\Omega) = A(f_0), \quad \text{i.e.} \quad A_{N_1,N_2}(f_0) \to A(f_0).$$

步骤 2. 令 $\tau := A^{-1}(r)$。任取 $\delta \in (0, \varepsilon_1)$。由于 A 严格递减且连续，

$$A(\tau + \delta) < A(\tau) = r < A(\tau - \delta).$$

定义正裕度

$$\eta := \frac{1}{4}\min\{\, r - A(\tau+\delta),\, A(\tau-\delta) - r \,\} > 0.$$

根据步骤 1，选取 M_1, M_2，使得对所有 $N_1 \geq M_1, N_2 \geq M_2$,

$$\left| A_{N_1,N_2}(\tau \pm \delta) - A(\tau \pm \delta) \right| < \eta.$$

Then

$$A(\tau+\delta) \leq r - 4\eta \quad \text{and} \quad A_{N_1,N_2}(\tau+\delta) < (r-4\eta) + \eta,$$

因此

$$A_{N_1,N_2}(\tau+\delta) \leq r - 2\eta.$$

同理,

$$A_{N_1,N_2}(\tau-\delta) \geq r + 2\eta.$$

接下来进一步增大 M_1, M_2，使得 $12/N_1N_2 < \min\{\varepsilon_2, \eta\}$。取 $c \in \mathbb{N}$ 为 rN_1N_2 的最近整数。于是

$$\left| r - \frac{c}{N_1N_2} \right| < \varepsilon_2 \quad \text{and} \quad r - \eta < \frac{c}{N_1N_2} < r + \eta,$$

因此特别地

$$A_{N_1,N_2}(\tau+\delta) < \frac{c}{N_1N_2} < A_{N_1,N_2}(\tau-\delta).$$

最后，利用 $A_{N_1,N_2}(\cdot)$ 是上水平集的归一化计数泛函这一解释。不等式$A_{N_1,N_2}(\tau+\delta) < cN_1N_2$意味着，值大于或等于 $\tau+\delta$ 的格点少于 c 个，因此第c大值小于*smaller than* $\tau+\delta$。同样，$A_{N_1,N_2}(\tau-\delta) > cN_1N_2$意味着，值大于或等于 $\tau-\delta$ 的格点多于 c 个，因此第c大值至多为*at most* $\tau-\delta$。因此

$$\tau - \delta \;\leq\; f_{*(N_1,N_2)}\Big(\frac{c}{N_1N_2}\Big) \;\leq\; \tau + \delta,$$

并回想 $\tau = A^{-1}(r)$,

$$\left| A^{-1}(r) - f_{*(N_1,N_2)}\Big(\frac{c}{N_1N_2}\Big) \right| \leq \delta < \varepsilon_1.$$

结合 $\left| r - cN_1N_2 \right| < \varepsilon_2$，证明至此完成。 □

命题 B.2（排序函数的闭式）。 固定 $t_1, t_2 > 0$。对于 $f_0 \in (0,1]$, 定义上水平集

$$\Omega(f_0) := \big\{ (x_1,x_2) \in [1,\infty)^2 : x_1^{-t_1} x_2^{-t_2} \geq f_0 \big\},$$

并令 $A(f_0) :=$ 面积 $\Omega(f_0)$。那么：

1. 如果 $t_1 \neq t_2$,

$$A(f_0) \;=\; 1 + \frac{t_1 f_0^{-1/t_1} - t_2 f_0^{-1/t_2}}{t_2 - t_1}.$$

2. 若 $t_1 = t_2 = t$,

$$A(f_0) \;=\; f_0^{-1/t}\Big(\tfrac{1}{t}\ln\tfrac{1}{f_0} - 1\Big) + 1.$$

此外，$A:(0,1]\to[1,\infty)$严格递减，因而是双射；连续秩频曲线是其反函数

$$f_*(r) \;=\; A^{-1}(r) \qquad (r \ge 1).$$

证明。由 $x_1^{-t_1}x_2^{-t_2} \ge f_0$ 和 $x_1,x_2 \ge 1$ 可得

$$x_1 \;\le\; f_0^{-1/t_1}\, x_2^{-t_2/t_1}.$$

因此，对于每个固定的$x_2 \ge 1$，可取的x_1区间为

$$x_1 \in \left[1,\; f_0^{-1/t_1} x_2^{-t_2/t_1}\right],$$

仅当$x_2 \le f_0^{-1/t_2}$时，才会产生正宽度。因此

$$\Omega(f_0) = \left\{(x_1,x_2):\; 1 \le x_2 \le f_0^{-1/t_2},\;\; 1 \le x_1 \le f_0^{-1/t_1} x_2^{-t_2/t_1}\right\},$$

and

$$A(f_0) = \int_1^{f_0^{-1/t_2}} \left(f_0^{-1/t_1} x_2^{-t_2/t_1} - 1\right) dx_2.$$

若$t_1 \ne t_2,$，记作 $\alpha := t_2/t_1 \ne 1,$

$$\begin{aligned} A(f_0) &= f_0^{-1/t_1}\, \frac{x_2^{1-\alpha}}{1-\alpha}\Big|_1^{f_0^{-1/t_2}} \;-\; \left(x_2\Big|_1^{f_0^{-1/t_2}}\right) \\ &= 1 + \frac{t_1\, f_0^{-1/t_1} - t_2\, f_0^{-1/t_2}}{t_2 - t_1}. \end{aligned}$$

若$t_1 = t_2 = t$，则上界为$f_0^{-1/t}\, x_2^{-1}$，因此

$$\begin{aligned} A(f_0) &= \int_1^{f_0^{-1/t}} \left(f_0^{-1/t} x_2^{-1} - 1\right) dx_2 \\ &= f_0^{-1/t}\, \ln\!\left(f_0^{-1/t}\right) - \left(f_0^{-1/t} - 1\right) = f_0^{-1/t}\left(\tfrac{1}{t}\ln\tfrac{1}{f_0} - 1\right) + 1. \end{aligned}$$

最后，对f_0中的任一表达式求导可得$A'(f_0) < 0$，因此A在$(0,1]$上严格递减，并存在定义良好的反函数A^{-1}，由此得到连续秩频曲线$f_*(r) = A^{-1}(r)$。 □

表3。经过归一化拟合的乐曲级齐夫–曼德布罗特参数。此处，“乐器”表示该乐曲中存在的乐器声部数量。上标 a 和 b 表示该乐曲所采用的时长归一化方案。标记为 a 的曲目族按 $\times 1.5$ 的因子进行缩放，而标记为 b 的曲目族按 $\times 0.5$ 的因子进行缩放。括号中的标签（G、J、P、Ch）表示语料库中的来源曲集。G 代表管乐灵山会相，J 代表中广之曲，P 代表平调会相，Ch 代表吹打。

曲名	乐器	q	s	N	L	R^2
别路打令[b]	4	34.804	3.697	79	1225	0.985
加乐打令（G）	4	288.958	20.000	47	426	0.987
加乐打令（J）	5	16.705	1.795	83	542	0.981
加乐打令（P）	5	48.642	3.895	56	560	0.976
길군악	4	14.821	2.695	63	1357	0.995
吉打令[b]	4	7.227	1.698	55	892	0.990
军乐（G）[b]	4	1.087	0.905	81	1355	0.936
军乐（J）[b]	5	4.869	1.298	112	1755	0.987
军乐（P）[b]	5	5.015	1.379	92	1726	0.979
军乐（Ch）[b]	4	1.069	0.935	83	1404	0.955
界面调数大叶——中石[a]	6	10.915	1.437	176	1506	0.985
界面二数大叶——言约[a]	6	10.520	1.321	213	1478	0.981
界面歌乐都德里[b]	5	4.630	1.464	94	1490	0.970
庆丰年[a]	4	0.872	0.923	131	1182	0.969
Hahyeondodeuri	5	3.287	1.200	107	898	0.960
Jungnyeongsan (G)	5	21.015	2.548	98	2021	0.985
Jungnyeongsan (J)	4	22.858	2.326	99	1291	0.991
Jungnyeongsan (P)	5	5.316	1.324	127	1888	0.988
慢拍正式	5	35.507	4.007	74	1846	0.989
Mitdodeuri[a]	5	6.188	1.853	103	2432	0.992
三弦道德里	4	5.582	1.512	70	887	0.947
上弦道德里（J）	5	2.955	1.254	103	1247	0.977
上弦道德里（P）	5	8.637	1.981	83	1284	0.990
上灵山（G）	4	8.173	1.369	117	1015	0.975
上灵山（J）	5	1.695	0.927	124	1877	0.946
上灵山（P）	5	10.512	1.611	121	2269	0.972
细灵山（G）	4	104.893	8.519	53	655	0.986
细灵山（J）	5	1.085	0.978	88	821	0.979
细灵山（P）	5	2.780	1.126	64	889	0.976
寿齐天	4	74.290	3.935	138	1319	0.990
水龙吟[a]	4	1.476	1.091	68	817	0.930
打令（G）[b]	4	9.326	2.050	51	888	0.941
打令（J）[b]	5	8.374	1.835	102	1493	0.983
打令（P）[b]	5	7.580	2.026	78	1461	0.992
羽调初数大叶 - 东唱[a]	6	8.328	1.230	170	1417	0.978
羽调加乐叟돌이[b]	5	15.068	2.308	89	1400	0.989
羽都드리[a]	5	6.193	2.002	89	2294	0.989
两清都드리乐章 7[b]	5	11.005	1.821	73	380	0.989
两清都드리乐章 1~6	5	22.335	3.963	71	1187	0.988
念佛都드리（G）	4	24.357	2.892	78	1262	0.983
念佛都드리（J）	5	20.145	2.270	122	1681	0.990
念佛都드리（P）	5	55.026	5.522	89	1672	0.993
与民乐乐章 1[a]	5	26.737	2.310	151	2963	0.986
与民乐乐章 2[a]	5	41.845	3.094	138	3229	0.986
与民乐乐章 3[a]	5	37.642	2.907	140	3125	0.989
Yeomyangchun[a]	4	19.320	1.846	123	1011	0.985

表 4. R^2 联合曲线模型中齐夫–曼德布罗特拟合的得分。行表示s，列表示t。

面板A：归一化拟合。

$s\backslash t$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
0.5	0.9990	0.9994	0.9997	0.9998	0.9999	0.9999	0.9999	1.0000
1	—	0.9986	0.9989	0.9993	0.9995	0.9997	0.9998	0.9998
1.5	—	—	0.9990	0.9993	0.9995	0.9996	0.9997	0.9998
2	—	—	—	0.9995	0.9996	0.9998	0.9998	0.9999
2.5	—	—	—	—	0.9998	0.9999	0.9999	0.9999
3	—	—	—	—	—	0.9999	0.9999	1.0000
3.5	—	—	—	—	—	—	1.0000	1.0000
4	—	—	—	—	—	—	—	1.0000

B面板：原始拟合。

$s\backslash t$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
0.5	0.9630	0.9954	0.9985	0.9992	0.9996	0.9997	0.9998	0.9998
1.0	—	0.9987	0.9991	0.9995	0.9996	0.9998	0.9998	0.9999
1.5	—	—	0.9993	0.9996	0.9997	0.9998	0.9999	0.9999
2.0	—	—	—	0.9997	0.9998	0.9999	0.9999	0.9999
2.5	—	—	—	—	0.9999	0.9999	1.0000	1.0000
3.0	—	—	—	—	—	1.0000	1.0000	1.0000
3.5	—	—	—	—	—	—	1.0000	1.0000
4.0	—	—	—	—	—	—	—	1.0000