

参赛学生姓名：_____高章涵_____

中学：_____中国人民大学附属中学_____

省份：_____北京市_____

国家/地区：_____中国_____

指导老师姓名：_____苏冠男_____

指导老师单位：_____北京中学_____

论文题目：_____基于群论与多目标优化的运动音乐生成模型_____

基于群论与多目标优化的运动音乐生成模型

高章涵

摘要：随着智能穿戴设备的普及和国家全民健身的深入开展，运动者通过音乐提升运动水平和持续时间的需求日益凸显。传统的静态音乐歌单难以适应运动过程中强度的动态变化，动态的多样化音乐选择能够更好地满足不同类型运动者的个性化需求。本文针对运动音乐的动态生成问题，基于群论和音乐生成之间关联的数学理论，提出一种基于群论与多目标优化的音乐生成模型。先给出一个简化模型建立框架，即把钢琴伴奏和弦中的大小三和弦建模成 24 阶阿贝尔群，并用样本乐曲统计音程变换，建立一致性与多样性两个评价函数。在此基础上，把节奏和音高建模到同一个直积群里，给出音乐模式成为子群的条件，基于最大公约数定义音乐复杂度，建立它和运动强度之间的映射关系，再给出实现复杂度平滑过渡的音乐变换模型，以及包含运动适配度、音乐连续性和听觉舒适度目标的动态权重综合评价。仿真结果表明，低运动强度时复杂度和目标比较接近，运动强度很高时优化可以灵活调整为保证连续性和舒适度目标。将本文与其它方法的结果比较，验证了本文理论模型的有效性。本研究为运动音乐生成提供了实用可验证的数学模型。

关键词：运动音乐生成；循环群；直积群；子群判定；音乐复杂度；多目标优化；动态权重

目录

1	研究背景.....	4
2	问题分析.....	6
2.1	文献和数学工具选择分析.....	6
2.2	研究目标.....	6
2.3	问题假设.....	7
2.4	符号定义.....	7
3	和弦进行的群论模型.....	8
3.1	十二平均律与大小三和弦.....	8
3.2	群和阿贝尔群.....	9
3.3	基于群论的和弦建模.....	9
3.4	音程变换的概率与评价函数.....	11
3.5	谱例数据分析.....	12
4	运动音乐的模型构建.....	13
4.1	节奏和旋律综合的音乐模式构建.....	14
4.2	运动适配的复杂度度量.....	18
4.3	使得音乐复杂度平滑过渡的变换.....	21
4.4	动态权重的多目标优化.....	23
5	仿真与结果分析.....	26
5.1	对于定理 2 子群满足条件的计算分析.....	27
5.2	对于定理 3 复杂度公式和平方映射的可达范围分析.....	27
5.3	平滑过渡变换的计算分析.....	28
5.4	动态加权选优的分析.....	29
5.5	连续运动过程及与其它模型的结果对比.....	30
6	结论.....	32
7	参考文献.....	34

1 研究背景

随着全民健身理念的普及和智能穿戴设备的广泛应用，个性化、智能化的运动健康服务需求日益增长。在各类运动场景中广泛证实，音乐能够有效调节运动节奏、提升耐力、延缓疲劳并增强运动愉悦感。因此音乐已经成为了人们运动中不可或缺的辅助工具。在运动过程中，所使用的音乐的节奏、速度甚至情感色彩如果能跟人体的运动状态（例如心率、步频、运动强度）相匹配，可以显著优化人们的运动表现和体验。然而，当前人们在运动时普遍依赖于静态、预先编制的音乐歌单。这类音乐缺乏对运动者生理状态和运动强度变化情况的感知和响应。例如，当运动者从热身阶段的低强度慢跑过渡到高强度的无氧间歇训练时，如果背景音乐的特征（如节奏密度、旋律复杂度）没有能随之进行平滑并且适配的调整，就不能发挥其积极的辅助作用，甚至可能因为节奏错位或情绪不符而产生干扰，影响运动的连贯性和效果。由此受到启发，如果能够研究开发实时感知运动状态并动态生成适配音乐的技术产品，对于提升运动科学性和个性化体验具有重要的意义和价值。

随着大数据和人工智能技术的发展，当前该方向研究已经取得一定进展，但面向运动场景的动态音乐生成研究仍处于初步阶段。现有的一些音乐生成方法有的是依赖于数据输入的预置规则库进行有限参数的切换的方法，这种灵活性差，难以覆盖复杂多变的运动状态；还有的是基于深度学习模型进行风格的模仿，这种情况系统往往作为一个“黑箱”，存在生成逻辑不透明、生成结果的可解释性与可控性较差的问题。运动场景这一应用要求所生成的音乐不仅是结构正确的，而更重要的是，生成过程必须要基于明确的、可量化的运动-音乐映射关系，以确保生成音乐能够匹配运动强度，或者避免听觉突兀。因此，探索一种数学理论上严谨的、能够清晰描述音乐结构及其与运动参数关联的“白箱”模型，是实现高质量的运动音乐动态生成的关键所在。

音乐中核心的节奏与音高体系，是具有周期性和循环性的。这为应用数学工具进行建模提供了天然基础。一方面，群论作为研究对称性与结构的数学分支，是描述此类音乐系统的有力工具。有国内外研究者在此方向上进行了持续的探索，但应用于我们所说的运动音乐生成场景也有一些可改进或完善的方面。文献[1]指出，十二平均律下的音高关系可以构成循环群。但该文献和现有其它研究中，大多局限于对音高或节奏等单一维度的群结构进行描述，缺乏将节奏与旋律统一起来进行建模的理论框架，因此难以支撑运动场景下节奏和旋律协同演变的需求。另一方面，为了实现对音乐的定

量化的感知和控制，需要建立有效的度量评价指标。有研究者利用数论中的一些概念，例如欧拉函数、最大公约数等，用来描述音阶生成元的性质[2]，但是并没有应用于构建综合的、可以与运动强度等外部变量动态关联的、评价音乐复杂程度等性质的模型。

另外，为了生成好的运动音乐，需要在多个相互冲突的目标之间进行权衡优化。我们考虑的目标包括：运动适配度（即音乐特征与当前运动强度的匹配程度）、音乐连续性（即音乐风格变化时能够进行平滑过渡）、听觉舒适度（即所生成的音乐本身是否和谐悦耳）。传统的多目标优化方法常使用固定的权重来加权多个目标，对于运动强度及其变化快慢的适应性较差。因此，需要设计一种根据运动状态变化自适应调整权重的动态优化机制。

综上所述，当前面向运动场景的动态音乐生成研究面临三方面挑战：第一，缺乏能够统一描述节奏-旋律二维结构及其协同演进的数学模型；第二，没有能够将音乐内在的特征转化为可评估的复杂度量化指标，也没有与运动强度之间建立可解释的映射关系；第三，没有建立能随着运动状态调整、动态权衡优化多个目标的决策框架。为此，本研究通过融合群论、数论与多目标优化等数学工具，构建运动音乐生成的理论模型和算法框架。

本文的架构如下。第 2 章进行问题分析，包括所选用的数学工具分析、研究目标、问题假设、符号体系等；第 3 章先对问题进行了一个简化研究，不考虑节奏和运动强度，只研究和弦进行的生成。在这里我们引入群论工具，把大小三和弦定义成有序数对，验证它们连同一种新运算构成阿贝尔群，这样就可以在这个群上进行相应运算。然后通过样本乐曲的音程变换进行统计，并利用两个简化的评价函数（一致性和多样性）进行评价，形成闭环。第三章的研究是基于本文作者的前期研究基础 [3]、面向本研究的需求进行的适应性修改。在这个基础上，第 4 章综合考虑节奏和音高，建模成一个直积群，并进行相应的变换分析和综合评估。第 5 章给出仿真与结果分析，第 6 章是本文结论。

2 问题分析

2.1 文献和数学工具选择分析

为了实现运动音乐的动态生成与适配，需要建立能够描述音乐结构和特征、并且支持优化决策的数学模型。经调研分析，现有的数学工具中，以下几类是与本问题相关的。

群论与模运算：音乐是基于音阶和节拍体系的，具有鲜明的周期性循环特征。循环群，特别是模整数加法群（例如用 Z_{12} 描述十二平均律的音高），为刻画音高、节奏的周期性提供了天然的数学语言 [1]。但是正如前文所说，现有研究多集中于单一的维度，缺乏将节奏与旋律统一建模的理论框架。

数论与欧拉函数：在音乐生成中，生成元的选择至关重要，它决定了音乐模式的覆盖范围和变化的丰富程度。欧拉函数 $\varphi(n)$ 给出与 n 互质的正整数个数，文献[2]利用这个性质来筛选能够生成完整音阶的生成元，实现了音乐生成元研究的定量化。但是其并没有考虑音乐复杂度的综合度量，或者与运动强度等外部变量建立联系。这为我们的研究中，评估音乐模式的“生成潜力”或“复杂程度”提供了研究思路。

多目标优化理论：运动音乐生成过程需要权衡多个目标，包括是否适配运动、听觉是否舒适、风格是否连贯等。例如，高强度的运动需要高复杂度音乐以提供激励，但过高的复杂度可能会降低听觉的舒适度。多目标优化理论提供了这类问题的求解框架。然而，运动音乐这一动态场景中，需要将权重根据实时运动状态（强度、变化率等）进行自适应调整，权重的分配调整策略需要进行特定设计。

2.2 研究目标

基于上述已有研究和所存在问题的分析，本研究的目标为：

第一，基于群论建立运动音乐的描述模型。基于循环群理论进行拓展，将节奏和旋律综合建模，形成综合性的音乐模式描述，确保所生成音乐符合实际应用条件，为后续工作提供理论基础。

第二，构建能够适配运动场景的音乐复杂度度量模型。定义可计算的音乐复杂度，并考虑与运动强度之间的内在联系，为实现音乐与运动之间的适配关联提供量化的依据。

第三，针对运动强度的动态变化带来的音乐形态切换突兀的问题，设计平滑的变换算法。需要遵循的原则是保证音乐模式切换过程中的复杂度守恒，从而实现音乐风格的自然过渡。

第四，提出权重可自适应动态变化的多目标优化框架。构建综合评价函数。设计动态的权重分配策略，求解出不同运动场景下的最优音乐模式。

2.3 问题假设

为了保证模型的有效性和可行性，本研究基于以下假设开展。

音乐结构的周期性假设：用于运动场景的背景音乐在时间（节奏）和音高（旋律）维度上均具有周期性循环的特征。这个假设是应用群论进行建模的前提。

复杂度可度量性假设：假设音乐模式的听觉复杂度，即丰富程度和多变性，可以通过其数学上的生成结构进行有效的量化。这个假设是建立复杂度度量模型并且与运动强度进行关联的基础。

运动-音乐映射假设：假设运动强度与运动者对音乐复杂度的需求之间存在确定的函数关系。实际情况下，高强度运动需要复杂多变的音乐激励，低强度运动则适合更简单重复的音乐来帮助放松。本文的模型据此进行设计。

听觉舒适度可量化假设：假设所听到的音乐是否舒适可以基于音程是否协和、节奏是否平稳、旋律是否流畅等可量化的特征进行评价。这个假设是定义听觉舒适度目标函数的依据。

实时性与简化性假设：假设运动音乐生成系统需要满足实时响应的要求，因此算法需要具备较低的计算复杂度。同时为了便于理论分析和初始验证，本研究忽略了复杂的音乐要素，例如复杂的和声或配器变化等，留作后续深入研究。

2.4 符号定义

本部分给出所使用的符号体系定义如下表。

表1 本文主要符号

符号	含义	数学定义 / 取值范围
T	三和弦集合	24个有序数对(r, m)，见定义2
(r, m)	一个三和弦	r为根音，m为+或-，见定义1
*	和弦运算	见(1)(2)式
E, H	一致性、多样性	见(4)(5)式，取值越大越好
N	节奏模数	时间分辨率，正整数（如N=8表示八分音符网格）
C	音高模数	音阶音级数，正整数（如C=12为十二平均律）

Z_N, Z_C	模整数加法群	元素为 0 到 $N-1$ (或 $C-1$)，运算为加法取模
f_r	节奏生成元	$1 < f_r < N-1$ ，控制相邻节奏事件的时间间隔
f_m	音高生成元	$1 < f_m < C-1$ ，控制相邻音高的音程大小
d	音乐片段长度	$d \leq \min(N, C)$ ，表示音事件数量
$M(f_r, f_m, d)$	音乐模式	$\{(kf_r \bmod N, kf_m \bmod C) k = 0, 1, \dots, d-1\}$
I	运动强度	$0 < I < 1$ ，实时采集的运动努力程度
$C(M)$	音乐复杂度	基于最大公约数的量化指标，取值 $(0, 1]$
$f_1(M)$	运动适配度	衡量音乐与运动强度的匹配度，取值 $[0, 1]$
$f_2(M)$	音乐连续性	衡量新旧音乐的风格相似度，取值 $[0, 1]$
$f_3(M)$	听觉舒适度	基于音程协和度的量化评分，取值 $[0, 1]$

3 和弦进行的群论模型

我在钢琴演奏的学习过程中，希望了解一些乐曲创作的原理，并学习一些即兴伴奏的知识，从而对所练习的钢琴曲有更深入的理解。即兴伴奏是指在事先毫无准备的情况下，临时根据主旋律配以丰富的和声和音型，边创作，边用乐器伴奏。即兴伴奏需要以左右手和声为基础，基于各种和弦与旋律相互融合，创造出好听并富有艺术感染力的音乐[4]。

本章面向即兴伴奏实际应用，研究一个简化问题，即旋律中的伴奏和弦的生成问题，即已经有一个和弦了，下一个和弦应该怎么走。在当前的科技发展进程中，已经有了一些智能软件工具，可以帮助人们进行乐曲即兴伴奏创作的自动生成。这些工具利用人工智能和机器学习技术，来分析输入的主旋律或和弦进行，然后生成适合的伴奏。但是，这类基于数据的生成过程往往缺少清楚的数学解释，限制了人们对生成过程的探究和改进。本章受到相关研究的启发[5][6][1]，利用抽象代数中的群来描述和弦及其转换关系，再讨论生成下一个和弦的概率，并建立评价函数来进行评估。

3.1 十二平均律与大小三和弦

将一个八度内的音程等比例地分成十二个等份，每一份称为一个半音，这就是十二平均律，是现代国际通用的标准音律。这个概念最早可以追溯到 16 世纪中国的朱载堉，他提出的“新法密率”实际上就是十二平均律的早期形式[5]。一个八度的频率范围被分成十二个等比的频段，每个半音的频率与其相邻半音的频率之间有一个固定的比率。

大三和弦和小三和弦是音乐中最常见的两类和弦。设一个三和弦的根音为 r ，则大三和弦的组成为 $(r, r+4, r+7)$ ，小三和弦的组成为 $(r, r+3, r+7)$ （音高都按十二平均律理解）。例如，在 C 大调中，C-E-G 就是一个大三和弦；Am (A-C-E) 就是一个小三和

弦。大三和弦通常给人一种明亮、欢快的感觉，而小三和弦则通常给人一种柔和、忧郁的感觉。这里的“大”和“小”并不是指音符的数量，而是指和弦中音程的性质。

3.2 群和阿贝尔群

由于和弦不仅仅是一组数的集合，还有大小和弦的特征区分，如果只用传统的整数集合和加减乘除，不利于表征这种复杂特性，也不方便进行和弦转换的运算建模。而群的概念是将数或其它不是数的元素组成集合，并增加自定义的运算而形成，便于形成统一的数学模型进行运算[6]。

群是一种代数结构，由一个集合以及该集合上定义的一个二元运算组成。这个运算需要满足四个基本性质，即封闭性、结合律、单位元存在和逆元存在。具体来说，群是一个集合 G ，以及在该集合上定义的一个二元运算 $*$ ，满足：

封闭性：对于 G 中的任意两个元素 a 和 b ，它们的乘积 $a*b$ 仍然在 G 中。

结合律：对于 G 中的任意三个元素 a 、 b 和 c ，都有 $(a*b)*c = a*(b*c)$ 。

单位元存在：存在一个元素 $e \in G$ ，使得对于 G 中的任意元素 a ，都有 $e*a = a*e = a$ 。元素 e 称为群的单位元。

逆元存在：对于 G 中的任意元素 a ，都存在一个元素 $b \in G$ ，使得 $a*b = b*a = e$ 。元素 b 称为元素 a 的逆元，记作 a^{-1} 。

如果一个群还满足交换律，即对于 G 中的任意两个元素 a 和 b ，都有 $a*b = b*a$ ，那么这个群就被称为可交换群，也叫阿贝尔群。

下面是几个例子。整数和加法是一个群，也是一个阿贝尔群。整数和减法不是一个群，因为不满足结合律；整数和乘法也不是一个群，因为单位元是 1，逆元不一定在整数集合里。对魔方的操作组成的魔方群是群，但不是阿贝尔群，因为对魔方上层转动一下然后左侧转动一下，结果跟左侧转动一下然后上层转动一下是不同的。矩阵和乘法组成一个群，但不是阿贝尔群，因为矩阵乘法一般不可交换。

3.3 基于群论的和弦建模

由于三和弦是音乐理论中最常见的和弦，而且在伴奏中其它类型的和弦大多可以用三和弦近似得到较好的和声效果，因此本研究中仅考虑三和弦。

定义 1：和弦用有序数对 (r, m) 表示，其中 r 为和弦根音， m 是大小和弦标记，用 $+$ 或 $-$ 表示。例如 $(0, +)$ 代表 C 大调大三和弦， $(0, -)$ 代表 C_m 小三和弦， $(8, -)$ 代表 $G\sharp$ 小调小三和弦。

表 2 三和弦标记、名称和有序数对表示（部分）

和弦标记	名称	(根音位置, 大小)
C	C 大调主和弦（大三和弦）	(0, +)
D	D 大调主和弦（大三和弦）	(2, +)
F	F 大调主和弦（大三和弦）	(5, +)
G	G 大调主和弦（大三和弦）	(7, +)
Cm	Cm 小调主和弦（小三和弦）	(0, -)
Dm	Dm 小调主和弦（小三和弦）	(2, -)
Am	Am 小调主和弦（小三和弦）	(9, -)

定义 2：以上全部 24 个和弦的有序数对组成的集合记为 T ，因此 T 中的元素个数为 24。

由于即兴伴奏的和弦生成需要考虑和弦演进的规律，从而对下一个出现的和弦进行预测，这就需要对有序数对之间进行运算。当前已有的运算规则无法对这样的有序数对直接运算，因此我们给出定义 3 如下。

定义 3：对 T 中任意两个元素 (r_1, m_1) 、 (r_2, m_2) ，规定运算

$$(r_1, m_1) * (r_2, m_2) = ((r_1 + r_2) \bmod 12, m_1 \cdot m_2) \quad (1)$$

其中 $\bmod$ 为模运算， $m_1 \cdot m_2$ 定义为与逻辑“与”运算类似的规则，即

$$(+) \cdot (+) = (+), \quad (+) \cdot (-) = (-), \quad (-) \cdot (+) = (-), \quad (-) \cdot (-) = (-) \quad (2)$$

定理 1（三和弦集合构成 24 阶阿贝尔群）

集合 T 和定义 3 中的运算 $*$ 构成一个阶数为 24 的阿贝尔群。

证明：需要满足封闭性、结合律、单位元存在、逆元存在四方面要求的代数结构，就是一个群。另外对于阿贝尔群还需要证明满足交换律。

封闭性：容易验证，对于集合 T 中的任意两个元素，对它们进行 $*$ 运算的结果仍然在 T 中。

结合律：容易验证，对于集合 T 中的任意三个元素 a 、 b 和 c ，都有 $(a * b) * c = a * (b * c)$ 。

群的单位元为元素 $(0, +)$ ，因为任意其它元素与 $(0, +)$ 进行运算，都等于那个元素本身。

群的元素 (r, m) 的逆元为 $((12 - r) \bmod 12, m)$ ，因其与原元素做运算得到单位元。

容易验证，群中的元素满足交换律。群中的元素个数为 24，所以阶数为 24。定理 1 证毕。

这样，所定义的集合 T 和新运算 $*$ 与通常意义上的整数运算规律类似，便于进行大规模运算。同时，又通过比整数群更丰富的群元素和新运算，将音乐和弦的特征都体现进来，比起单纯通过整数进行建模，是更为精确和完整的数学模型。

定义 4: 集合 T 中的元素（即和弦）从 $(r1,m1)$ 到 $(r2,m2)$ 的音程变换为

$$((r2-r1) \bmod 12, m1 \cdot m2) \quad (3)$$

音程变换的个数与群的元素个数一致，也是 24 个。

3.4 音程变换的概率与评价函数

即兴伴奏的和弦生成需要用到统计和概率。统计是整理和分析数据，帮助我们了解数据的特征和关系；概率是描述一件事情发生的可能性的。在即兴伴奏智能生成中，首先要通过统计已有样本乐曲的和弦进行（在本研究的理论模型中叫做音程变换），得到哪些和弦进行出现的次数多、哪些出现的次数少；然后将这个统计结果作为生成下一个和弦的概率。

从音乐创作的角度看，智能生成的音乐有两方面评价标准需要考虑。一方面，一首乐曲的和弦进行要符合一定的规律，比如通常从主和弦开始，经历多种下属和弦或属和弦的排列组合，最终以主和弦结束。这体现在所生成的下一个和弦要与大多数乐曲的和弦进行相一致。这个评价标准称为“一致性”。另一方面，音乐创作需要不断演进，需要有与传统和弦进行不完全一致的情况，来体现变化多样。这个评价标准称为“多样性”。和弦的生成需要在这两个评价标准之间进行权衡。

基于若干首乐曲样本，统计得到 24 个音程变换每一个出现的比率为 s_i ($i = 0, \dots, 23$)，满足 $0 \leq s_i \leq 1$ 且它们的和为 1。智能生成时，每一个音程变换都有一个概率 p_i ，同样满足 $0 \leq p_i \leq 1$ 且它们的和为 1。 p_i 就是需要设计的变量。

定义 5 一致性函数

$$E = \sum_{i=0}^{23} s_i \cdot p_i \quad (4)$$

定义 6 多样性函数

$$H = -\sum_{i=0}^{23} p_i \cdot \log_2 p_i \quad (5)$$

根据一致性评价标准，各 p_i 与对应的 s_i 越一致，一致性越好，这时 E 越大。根据多样性评价标准，各 p_i 越均衡，多样性越好，这时 H 越大。

用一个简单例子说明。假设样本里只有两个音程变换，出现比率分别为 0.2 和 0.8。如果完全按照一致性，生成概率也取 0.2 和 0.8，则 $E=0.68$ ；如果完全按照多样性，生

成概率取 0.5 和 0.5，则 $E=0.5$ 。可见 E 越大，一致性越优。反过来，第一种情况下多样性函数较小（约为 0.50），第二种情况下多样性函数较大（约为 0.69）。可见 H 越大，多样性越优。

实际生成需要同时照顾一致性和多样性，可以为 E 和 H 分别设置权重 w_1 和 w_2 ，建立综合目标 $w_1 \cdot E + w_2 \cdot H$ 。由于对该模型的求解涉及到优化方法和微积分知识，本章暂不往下求解，只基于统计数据进行分析。第 4 章的 4.4 节中，把同样的加权思想用到三个可以实时计算的目标上，是本章内容的扩展。

3.5 谱例数据分析

研究中选取了《虫儿飞》和《一笑江湖》两首乐曲。《虫儿飞》是经典传统的儿童歌曲，其和弦进行具有代表性；《一笑江湖》是近年流行的歌曲，乐曲变化相对多一些。对每一首谱例，先根据右手旋律和左手伴奏织体得到左手代表的和弦，再转化成有序数对，按(3)式计算音程变换，最后统计每个音程变换出现的比率。

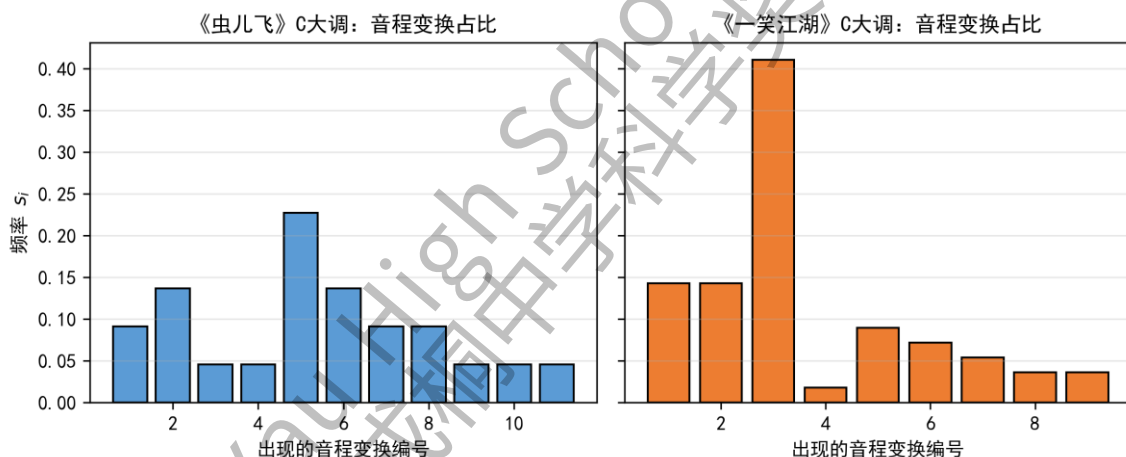


图1 《虫儿飞》《一笑江湖》C大调谱例中音程变换的占比

我们分析了互联网上公开的《虫儿飞》乐曲的四个不同版本的钢琴谱例（详见附录），其中用到的音程变换数量均在 10 个左右，占有最多比率的音程变换占比约为 20%，最少的约为 5%。占比较高的音程变换主要是在 F 大三和弦、G 大三和弦、C 大三和弦、Dm 小三和弦之间的转换。由此可以得出：在即兴伴奏智能生成过程中，为保证一致性，这几类和弦需要占主导作用；同时其它和弦的比率相对均衡，它们之间以相近概率出现可以提升多样性。

《一笑江湖》的三个谱例（详见附录）中，用到的音程变换数量为 8~9 个。C 大调谱例中占有最多比率的音程变换占比达到了 40%以上，对照乐谱可以发现这是由于乐曲中大量使用了从 Dm 小三和弦到 G 大三和弦的转换。

进一步计算评价函数。对《虫儿飞》C 大调谱例，当 p_i 取成和统计比率 s_i 一样时，一致性最好， $E \approx 0.124$ ，此时 $H \approx 3.227$ ；当 24 个音程变换等概率出现时，多样性最好， $H \approx 4.585$ ，此时 $E \approx 0.042$ 。对《一笑江湖》C 大调谱例，一致性最好时 $E \approx 0.228$ ， $H \approx 2.586$ 。两组谱例比较时，《一笑江湖》的一致性函数值更大，这和它的音程变换分布更加不均衡是一致的，说明所建立的评价函数和直观的数据分析结果相符。

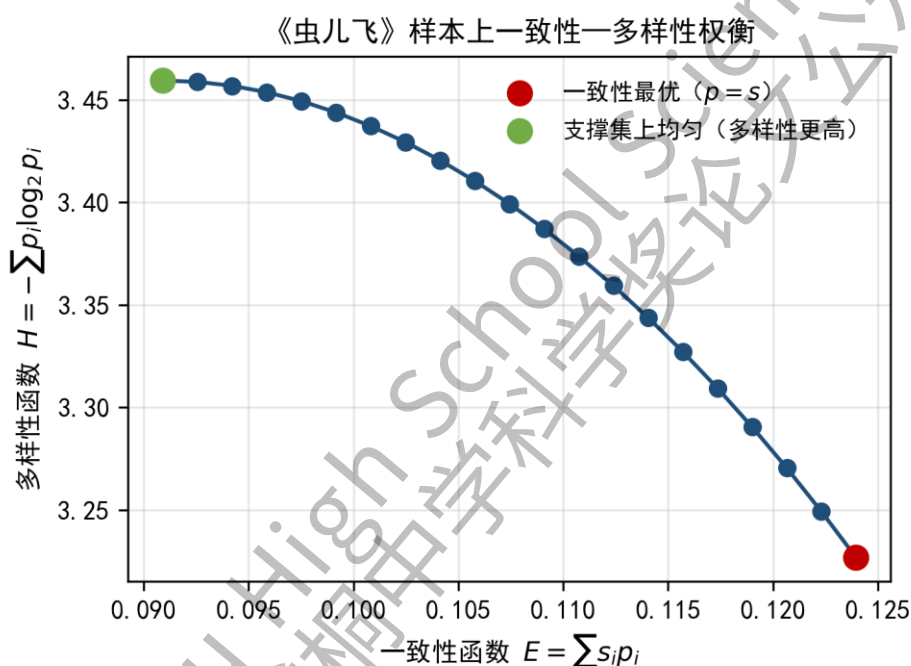


图2 《虫儿飞》样本上一致性函数 E 与多样性函数 H 的关系

通过本章研究，得到以下成果。第一，音乐对象可以写成群里的元素，对象之间的迁移变换可以通过群里的运算表达。第二，可以把两个相互矛盾的的评价进行分别打分，然后再加权。这就为后续章节中的进一步建模分析打下基础。

4 运动音乐的模型构建

这一章考虑运动场景的音乐。实际情况下，需要把节奏加入到建模考虑中，而且要能随着运动的强度而适应性的相对平滑变化。因此本章把这个问题拆开具体研究。

4.1 节奏和旋律综合的音乐模式构建

第 3 章用有序数对 (r, m) 表示一个和弦的方法，表示的是旋律中和声进行上的一个时间点。而真实场景下的音乐，包括本研究考虑的运动场景，都是同时包括旋律和节奏，而且旋律中除了和弦伴奏还需要考虑旋律线条的变化。旋律和节奏都有循环性的特点：一个小节里的节奏拍点会在后续小节中循环出现，十二平均律或其它类型里的旋律音高也会循环。沿着第 3 章的数学工具，在群论里，循环群是适合建模带有循环特征对象的工具，而直积群是可以将旋律和节奏合在一起进行处理的工具[1]。接下来首先简要介绍怎样从问题出发选择了这两种数学工具，然后介绍它们在本文问题背景下的定义。

4.1.1 为什么用循环群和直积群来表达音乐模式

考虑到例如运动过程的应用场景，跑步、跳绳时的背景音乐的是基于某种基本音型进行在作曲规则范围内的一小段动机重复。这要求时间格子和旋律线条都在本曲目的限定规则范围内进行可变的重复：节奏点在时间格子里特定的时间点出现，旋律的音高也是在所限定的范围内按规律出现的，即使升八度后的音名也相同。这就是引入循环群的原因。而又要把节奏和旋律写到一个对象里进行整体操作，而运算规则分开，这就是引入直积群的原因。这是本研究借鉴文献[1]所使用的建模方法思想。

4.1.2 循环群和直积群

设 n 是大于 1 的正整数，把 $0, 1, 2, \dots, n-1$ 这些数放在一个集合里，规定运算是普通加法后再除以 n 取余数（mod 运算）。这个集合连同这个运算记作 Z_n 。例如 $n=12$ 时，11 再加 1 得到 12，12 除以 12 余 0，所以 $11+1$ 就是 0，这就跟十二平均律里升十二个半音回到同一个音名相对应了。

按照群的性质进行检查：两个余数相加再取模，结果还在 0 到 $n-1$ 里，即满足封闭性；先加任意两个再加第三个，和先加不同的两个再加最后一个，余数是一样的，即满足结合律；加上 0 不变，即单位元是 0； a 的逆元是 $n-a$ （因为 $a+(n-a)$ 的余数是 0），0 的逆元仍是 0。所以 Z_n 是一个群。它同时还满足交换律，因而是阿贝尔群。这样构成的群还是循环群，是因为从单位元 0 开始，不断叠加生成元 1， n 次叠加生成元 1 后回到单位元 0，中间把所有元素都遍历了一遍。

注意：以上两段的符号 n 仅是为了介绍概念方便而举例，与下述的模型的数学描述不一样。接下来的符号中的 N 与节奏相关， C 与旋律相关，即跟 2.4 节的符号定义含义一致。

节奏方面，把一个小节均匀分成 N 个时间格子，格子的编号是 $0, 1, \dots, N-1$ ，一个小节走完 N ，到下一个小节再回到 0 ，以此类推，这就形成了一个循环群 Z_N 。类似地，旋律方面，把一种音律中的所有的旋律音高放在一个集合里，音高的数量即集合元素数量 C （如十二平均律 $C=12$ ），并且包含相应的运算，这样的循环群表示为 Z_C 。

这样，音乐中每一个发出的音，都可以写成一对整数 (t, p) ： t 是节奏的位置， p 是旋律音高。所有这样的数对放在一起，就构成了直积群，记作

$$Z_N \times Z_C = \{(t, p) | t = 0, 1, \dots, N-1; p = 0, 1, \dots, C-1\}$$

规定运算为两个分量各自做取模加法：

$$(t_1, p_1) + (t_2, p_2) = ((t_1 + t_2) \bmod N, (p_1 + p_2) \bmod C)$$

群的单位元是 $(0, 0)$ ； (t, p) 的逆元是 $((N-t) \bmod N, (C-p) \bmod C)$ 。例如，节奏取八分音符为单位即 $N=8$ ，音高遵照十二平均律即 $C=12$ 时，一共有 $8 \times 12 = 96$ 个格子点，这就构成了一段音乐所能走到的所有节奏和旋律组合的集合。

接下来，需要决定一段有限长的音型怎么从直积群表达的模型里取出来。音乐从所有格子点中的某个起点出发，按固定步长走 d 步，得到 d 个音，就得到了一小段音乐的模块。

设定 f_r 是相邻两个音之间隔了几个时间格子， f_m 是音高跳了几个单位音，分别作为各自的生成元，写成一对即是 (f_r, f_m) ：从 $(0, 0)$ 出发，第 k 个音就是把这个步长加 k 次。由此即可定义音乐模式如下。

定义 7 音乐模式

$$M(f_r, f_m, d) = \{(kf_r \bmod N, kf_m \bmod C) | k = 0, 1, \dots, d-1\} \quad (6)$$

例如，取 $N=8$ 、 $C=12$ 、 $d=4$ （一段里发 4 个音）、步长 $(f_r, f_m) = (4, 3)$ ，按(6)式进行计算：

$k=0$ 时，得到 $(0, 0)$ ：第一个音在小节开头、音高 0；

$k=1$ 时，得到 $(4, 3)$ ：时间加 4，音高加 3；

$k=2$ 时, 得到 $(8 \bmod 8, 6)=(0, 6)$, 因为时间 8 已经超出一小节, 余数是 0, 所以回到格子 0, 音高继续加到 6;

$k=3$ 时, 得到 $(4, 9)$ 。

如果再走一步 $k=4$, 就可以得到 $(8 \bmod 8, 12 \bmod 12)=(0, 0)$, 正好可以回到开头, 符合定义要求, 所以这 4 个音可以作为音乐模式使用。

以上例子如图 3 所示。

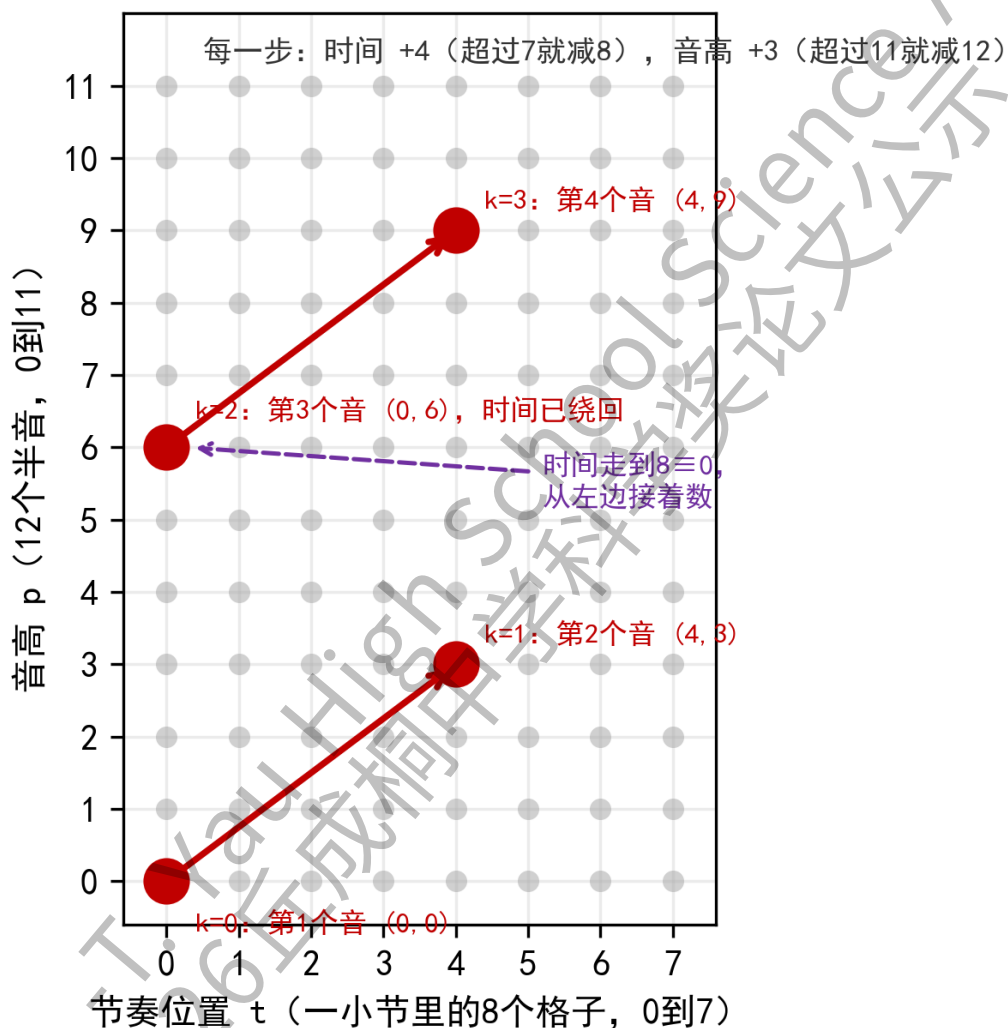


图 3 一段 4 个音的运动音乐模式。

图 3 中, 横轴代表一小节里的 8 个时间格子, 纵轴代表 12 个半音。浅灰点就是直积群里全部可能的(时间,音高)组合, 即 (t,p) 。红色的点是这段音乐模式里实际发出的音。

4.1.3 音乐模式生成所要遵循的规律

上一小节所举的例子中， $k=4$ 时恰巧可以回到 $(0, 0)$ 点，这样的音乐模式才能够使用，否则到下一轮循环时会错拍，生成的音乐就会乱。如何才能保证符合这样条件的模式？对此进行探究，这就需要满足：把步长 (f_r, f_m) 加 d 次后，能够回到等于 $(0,0)$ 。也就是说， $d \cdot f_r$ 是 N 的倍数的同时 $d \cdot f_m$ 是 C 的倍数。

希望音乐模式 M 能够满足群的其它性质，即满足封闭、结合、单位元、逆元等特征，以便于能够在群的框架下进行运算。而音乐模式 M 又是基于直积群 $Z_N \times Z_C$ 得出的，所以如果能判定音乐模式 M 是直积群 $Z_N \times Z_C$ 的子群，它就符合群的性质。因此有以下定理 2。

定理 2 需要解决以下问题，即什么时候由(6)式走出来的这 d 个点，刚好构成直积群的一个子群。条件找到以后，后面的平滑变换和优化都只在这些“合格模式”里选，生成出来的音才能按小节转圈。

定理 2（音乐模式构成直积群的子群）

音乐模式 $M(f_r, f_m, d)$ 构成直积群 $Z_N \times Z_C$ 的子群，当且仅当

$$d \cdot f_r \equiv 0(\text{mod} N) \quad \text{且} \quad d \cdot f_m \equiv 0(\text{mod} C) \quad (7)$$

此时 $M(f_r, f_m, d)$ 具有完美的周期性，适合作为运动音乐的基础模块。

证明：要证明 M 是 $Z_N \times Z_C$ 的子群，需满足子群判定的 3 个条件：①非空；②封闭性（群内任意两元素运算后仍在群内）；③逆元存在（群内任意元素的逆元仍在群内）。结合音乐模式的定义，分充分性和必要性两步证明。

充分性证明。若(7)式成立，则 M 是子群。

①非空性：当 $k=0$ 时， $(0 \cdot f_r \text{ mod } N, 0 \cdot f_m \text{ mod } C) = (0,0)$ 属于 M ，故 M 非空。

②封闭性：任取 M 中两个元素 $(k_1 f_r \text{ mod } N, k_1 f_m \text{ mod } C)$ 和 $(k_2 f_r \text{ mod } N, k_2 f_m \text{ mod } C)$ ，其中 $0 \leq k_1, k_2 \leq d-1$ 。两者在直积群中的运算为分量上的模加法。令 $k=k_1+k_2$ 。若 $k < d$ ，则结果仍在 M 中；若 $k \geq d$ ，由(7)式可知 $d f_r = tN$ 、 $d f_m = sC$ （ t, s 为正整数），此时可写成 $k = d + k'$ （ $0 \leq k' < d$ ），于是 $k f_r \equiv k' f_r (\text{mod } N)$ ，同理，音高分量也回到 M 中，满足封闭性。

③逆元存在性：任取 M 中元素 $(k f_r \text{ mod } N, k f_m \text{ mod } C)$ ，其逆元为 $(-k f_r \text{ mod } N, -k f_m \text{ mod } C)$ 。由 $d f_r \equiv 0(\text{mod } N)$ 可知 $(d - k) f_r \equiv -k f_r (\text{mod } N)$ ，同理音高分量也成立。因 $0 \leq d - k \leq d - 1$ ，故逆元仍在 M 中。

综上，充分性得证。

必要性证明。若 M 是子群，则(7)式成立。因为 M 是子群，且 $(f_r \bmod N, f_m \bmod C)$ 属于 M ，对它进行 d 次加法运算（封闭性），可得 $(df_r \bmod N, df_m \bmod C)$ 属于 M 。又因 M 中元素的第一个分量为 $0, f_r, 2f_r, \dots, (d-1)f_r \pmod{N}$ ，共 d 个不同元素（生成元定义），所以 $df_r \bmod N$ 只能为 0 （否则会出现重复元素，与生成元特性矛盾），即 $d \cdot f_r \equiv 0 \pmod{N}$ ；同理，也可以证得 $d \cdot f_m \equiv 0 \pmod{C}$ 。

综上，定理 2 的充分必要条件成立，证明完毕。

通过本节的内容，可以得到，无论怎样选择时间格子和旋律，只要依据所得到的定理 2 的条件进行构建，就能够生成可以成段循环的音乐。这样，真实的运动音乐使用场景中，可以据此构建不依赖于惯用的节奏型或十二平均律的、多种变化丰富的音乐，并将其有机组合，而其内在结构限定了其不会产生旋律或节奏的配合紊乱的错误。

至此，本节的核心内容叙述完毕。最后再增加一些跟现有其它模型之间的对比。对于音乐模式的生成，还可以用线性生成模型或随机生成模型等方法[1]，但是线性模型生成的音乐较为单一，随机生成就可能较难控制随机数来满足音乐的规整程度要求。因此本文提出的方法，在理论严谨的前提下，能够满足运动音乐的周期性和节律同步的需求。

4.2 运动适配的复杂度度量

运动场景中的音乐，还应该满足适配的问题，即运动强度高的时候，音乐元素的变化应该更丰富一些，以达到较好的效果；强度低的时候，可以使用慢频率变化的简单重复音乐，使人放松。也就是需要有一个面向运动适配程度的复杂度度量。所考虑的对象应该是音乐模式生成中的核心要素 f_r 和 f_m 。接下来研究这个问题。

运动时的强度是不断变化的，音乐需要能随着运动强度的变化而变化。对于跟运动强度挂钩的音乐复杂度，需要符合一些条件。一是要能快速实时计算，因为从运动者的可穿戴设备上感知得到其强度、并用于音乐生成调整的过程，不应当有特别大的运算量，更不应该做基于大量数据的复杂统计；二是含义要清楚直接，例如大数代表变化多，小数代表变化少。

而由 4.1 节的研究得出，由符合条件的步长 (f_r, f_m) 所生成的音乐，就有了确定的将走过哪些时间格子（节奏）、和哪些不同的音高（旋律）。音乐的复杂程度就应当刻画所生成的音乐的时间格子有多长、音高丰富程度有多少。

本研究的设计基于文献[2]中的思想启发。文献[2]指出，音阶生成元的数量与欧拉函数相关：与模数 n 互质的那些数，在循环群 Z_n 里能够走遍全部 n 个位置，从而可以生成完整的音阶。例如，当 $n=8$ 时，1、3、5、7 都和 8 互质， $\varphi(8)=4$ ，用它们当作节奏步长就可以遍历到一小节里全部 8 个格子；这就对应到我们所期望的最丰富、最不重复的音乐模式。所以这给我们的启发就是：生成元和模数是否互质，决定了这段音能不能遍历到所有的元素，因此就可以用来衡量所生成的音乐的丰富程度。

这样生成元和模数互质的情况，在我们的研究中就成了一种理想情况，是衡量丰富程度的标杆。由于我们不是像[2]中那样以生成完整音阶为目的，因此在二者不互质时，也就是在一个周期里没有遍历所有音的情况，满足条件的生成元也能使用，

我们的设计思路受启发于循环群理论里的一条常用的性质：从 0 出发每次加 a ，在模 n 下能走到的不同位置一共有 $n/\gcd(a,n)$ 个。 $\gcd(a,n)=1$ 时正好走满 n 个，也就是上面提到的用欧拉函数筛选出来的那批生成元的情况。随着 $\gcd$ 越大，能够覆盖到的位置越少，节奏或旋律越“简单重复”。所以，就可以使用这条性质，来设计函数，衡量所生成音乐模式的复杂程度。

下面通过定理 3 给出具体的评价方法。

定理 3（运动适配的音乐复杂度度量）

音乐模式 $M(f_r, f_m, d)$ 的复杂度为

$$C(M) = 1/[\gcd(f_r, N) \cdot \gcd(f_m, C)] \quad (8)$$

且运动强度 I 与目标复杂度的映射关系为平方映射

$$C_{target} = I^2 \quad (9)$$

证明：基于前述，音乐模式中节奏、音高的丰富度与多变性越高，所生成的音乐的复杂度越高。结合生成元控制节奏间隔与音高变化的作用，复杂度与生成元的“生成能力”正相关，而生成元的生成能力由其与模数的最大公约数决定。具体分析和推导如下。

1. 节奏维度复杂度推导：节奏生成元 f_r 在循环群 Z_N 中生成的子群规模为 $N/\gcd(f_r, N)$ （根据循环群中生成元 a 生成的子群阶为 $n/\gcd(a,n)$ 这一性质）。子群规模越大，节奏复杂度越高，因此节奏维度复杂度可表示为 $1/\gcd(f_r, N)$ （归一化后，取值范围在 0 到 1 之间）。

2. 音高维度复杂度推导：与节奏维度同理，音高维度复杂度为 $1/\gcd(f_m, C)$ 。

3. 整体复杂度合成：音乐复杂度是节奏复杂度与音高复杂度的协同作用，两者相互独立，所以可以将两者的乘积作为整体的复杂度。而且乘积可以保证整体取值范围仍然在 0 到 1 之间，而且在两者均高的情况下，整体也高，符合实际的情况，即(8)式。

4. 运动强度与目标复杂度的平方映射：结合运动场景实际需求，高强度运动（ I 接近 1）时，人体对音乐多变性的需求呈加速增长（需更丰富的节奏、音高激励运动）；低强度运动（ I 接近 0）时，需求增长平缓（需简单重复的音乐放松）。因此我们使用平方函数 $C_{target} = I^2$ 来建模这一规律。

由此得到定理 3 的结论。证明完毕。

以上的建模方法对本研究所述的应用场景是合适的。原因分析如下。第一，复杂度是基于 4.1 节模型中的 f_r 、 f_m 、 N 、 C 算出来的，没有额外增加新的变量。第二，基于运动场景音乐生成的实时性要求，这里的计算是整数运算，求 gcd 就能得到，便于快速得出结果。第三，函数的设计跟运动直觉一致，当 gcd 小的时候，所生成的音乐丰富，适合高强度；gcd 大的时候，音乐相对简单，适合放松。

作为例子，图 4 把平方映射的(9)式画出，同时计算了 $N=8$ 、 $C=12$ 、 $d=4$ 时得到的复杂度范围（绿色条带），可以看出这一特定情况下的复杂度取值跟理论曲线之间的关系。

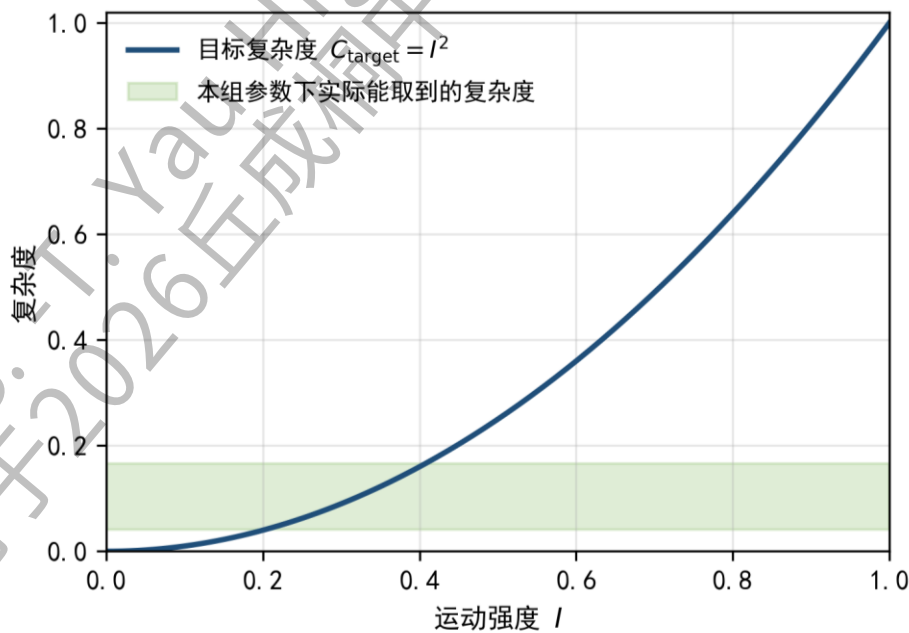


图 4 运动强度与目标复杂度的平方映射和示例的复杂度范围

（绿色条带为取值 $N=8$ 、 $C=12$ 、 $d=4$ 时得到的复杂度范围）

4.3 使得音乐复杂度平滑过渡的变换

基于运动强度的变化，音乐复杂度应该随之变化。但如果运动者的运动强度突然变化，设备感知到以后将会映射到音乐复杂度的变化指令，就容易导致乐段之间的切换突兀，节奏断层等问题，影响听觉感受。

因此，需要研究使得音乐复杂度平滑过渡的变换。这样，一方面在运动强度变化时可以以较为缓和的方式过渡到新的状态，又可以在需要保持当前复杂度不变的时候通过相同复杂度的变换来增加丰富程度。

由(8)式可知，复杂度由两个最大公约数决定。因此如果新的段落跟旧的段落最大公约数相同，就得到相同的 $C(M)$ 。设上一段音乐模式为 M_{prev} ，当前待生成音乐模式为 M_{curr} ，两者需满足

$$\gcd(f_r^{curr}, N) = \gcd(f_r^{prev}, N), \quad \gcd(f_m^{curr}, C) = \gcd(f_m^{prev}, C) \quad (10)$$

举例说明，采用跟图 3 一样的参数 $N=8$ 、 $C=12$ ，如果上一段取 $(f_r, f_m) = (4, 3)$ ，则 $\gcd(4, 8)=4$ 、 $\gcd(3, 12)=3$ ，复杂度 $C=1/(4 \times 3)=1/12$ 。如果下一段改成 $(4, 9)$ ， $\gcd(4, 8)$ 仍然是 4、 $\gcd(9, 12)$ 仍然是 3，这样后一段跟前一段相比复杂度不变，(10)式成立。但是如果下一段改成 $(2, 3)$ ，则得到 $\gcd(2, 8)=2$ ，复杂度就变成了 $1/6$ ，这样的旋律就会突然变得密集，听起来就会很突兀。

有了(10)式的限定，就可以选择合适的生成元，以满足此规则来保证音乐复杂度变换的平滑。

具体到某个音的变换，我们设计通过以下缩放加平移的方法进行。

第一步缩放，是把每个音的时间坐标乘一个整数 k_r 、音高坐标乘一个整数 k_m ，再各自取模。即进行如下操作：

$$t_{scale} = k_r \cdot t_{prev} \bmod N, \quad p_{scale} = k_m \cdot p_{prev} \bmod C \quad (11)$$

k_r 、 k_m 的取值要求 $\gcd(k_r, N) = 1$ 、 $\gcd(k_m, C) = 1$ ，也就是需要和格子的总数互质。否则会把不同的点压到同一个格点上。举例来说，当 $N=8$ 时，如果取 $k_r = 2$ ，原来的 $t=0$ 和 $t=4$ 乘 2 再模 8 就都变成了 0，两个音并成了一个，循环结构就被破坏了。而当互质时，例如 $N=8$ 时可选 1、3、5、7，这时候乘法就只是把 8 个时间格子重新排了一遍，点的数量没有变。

当运动强度提升时，可以适当增大 k_r 、 k_m ，同时仍须保持互质的条件，使节奏间隔更紧凑、音高变化更明显；强度降低时则取较小的 k_r 、 k_m ，使得节奏舒缓一些。

第二步平移，是把缩放后的所有点整体平移，时间增加 b_r 、音高增加 b_m 。这样用来跟前一段对齐避免错位。记为：

$$t_{curr} = t_{scale} + b_r \bmod N, \quad p_{curr} = p_{scale} + b_m \bmod C \quad (12)$$

b_r 可以取 0 到 $N-1$ ， b_m 可以取 0 到 $C-1$ 。例如，如果希望第一个音仍然在原来的时间、原来的音高上，那就取 $b_r = 0$ 、 $b_m = 0$ ；再比如说如果跑步时的节奏点总是卡不上，那就把 b_r 取成 1 或 2，整段往后挪。

把第一步的(11)和第二步的(12)合并，即进行如下操作：

$$M_{curr} = (k_r t_{prev} + b_r \bmod N, k_m p_{prev} + b_m \bmod C) \mid (t_{prev}, p_{prev}) \in M_{prev} \quad (13)$$

下面举例说明(13)的计算方法。还是采用图 3 的四个点(0,0)、(4,3)、(0,6)、(4,9)，如果取 $k_r = 5$ 、 $k_m = 7$ （都满足互质条件）、 $b_r = 1$ 、 $b_m = 2$ ，则

$$(0,0) \rightarrow (5 \cdot 0 + 1, 7 \cdot 0 + 2) \equiv (1, 2);$$

$$(4,3) \rightarrow (5 \cdot 4 + 1, 7 \cdot 3 + 2) = (21, 23) \equiv (5, 11) \quad (\text{因为 } 21 = 2 \times 8 + 5, 23 = 1 \times 12 + 11);$$

$$(0,6) \rightarrow (1, 7 \cdot 6 + 2 = 44) \equiv (1, 8);$$

$$(4,9) \rightarrow (21, 7 \cdot 9 + 2 = 65) \equiv (5, 5)。$$

这样，所得到的四个点都仍然在还在 8×12 的格子里。如图 5 所示，蓝色圆点是变换前的四个音，红色方点是按(13)式变换后的四个音，绿色箭头表示变换前后音的映射。这样就实现了复杂度平滑过渡的变换。

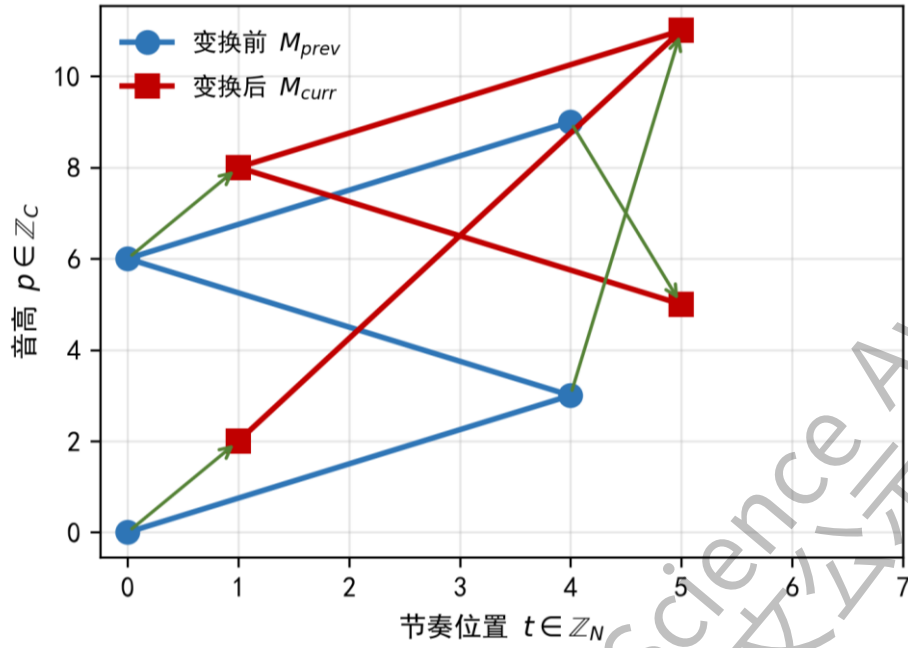


图 5 复杂度平滑过渡的变换示例

基于此，复杂度平滑过渡的算法执行步骤如下。

步骤 1：确定上一段音乐模式 M_{prev} 的参数，提取 f_r^{prev} 、 f_m^{prev} ，计算相应的最大公约数。

步骤 2：根据当前运动强度，确定待生成音乐模式的生成元，确保满足复杂度平滑过渡的条件(10)。

步骤 3：求解变换参数。缩放系数 k_r 、 k_m 分别从与 N 、 C 互质的整数中选择；平移系数 b_r 、 b_m 按实际步频的需要进行对齐。

步骤 4：将参数代入(13)式，将 M_{prev} 的每个元素进行变换，生成 M_{curr} 。

步骤 5：验证 M_{curr} 满足定理 2 的要求，验证通过后输出，完成变换。

4.4 动态权重的多目标优化

根据以上几节的研究结果，可以得到满足定理 2 和(10)式的生成元。然而这样符合条件的生成元往往不止一组。例如当 $N=8$ 、 $C=12$ 、 $d=4$ 时，就有 9 组合格模式。我们还需要根据运动场景的实际情况，进行优化以选择最好的。本节就分析这个问题。

根据第 2 章中我们设定的研究假设，运动场景里至少需要满足三方面要求。首先音乐要跟当前的运动强度符合，还有前后音乐段落要平滑过渡，再就是听觉感受要好。这三个方面是应该相互权衡的。例如高强度的运动需要更密集、变化更频繁的音，但

这样的音乐可能不太好听；如果复杂度的变化总是不大，那又可能跟不上运动强度的变化。

现有文献未涉及运动场景下的音乐选择的多目标协同优化问题。本部分研究对此进行探索。我们借鉴第 3.4 节里简化问题中的优化思路，考虑三方面的加权优化，把这三个方面先计算出优化目标值，然后进行加权求和。所改进的是，第 3 章的权重是固定值，我们这里将权重设计为受到运动强度 I 以及强度变化快慢 v 影响的函数，最后选择综合得分最高的那一组设计值。

下面我们给出具体三个目标函数的设计及综合得分计算方法。规定三个目标函数的取值都是 0 到 1 之间，越接近 1 越好，越接近 0 越差。下面介绍过程中的例子是沿用图 3 和图 5 里的模式 $M(4,3,4)$ 的情况。

①运动适配度 $f_1(M)$ ：考虑当前音乐的复杂程度和现在的运动强度是否匹配。由定理 3，当前强度 I 对应的目标复杂度是 $C_{target} = I^2$ ；这段模式自己的复杂度 $C(M)$ 由(8)式计算。两者越接近，匹配程度越好。设计运动适配度函数 $f_1(M)$ 如下：

$$f_1(M) = 1 - |C(M) - C_{target}|/\varepsilon \quad (14)$$

其中 $\varepsilon=0.05$ 是允许误差。如果分子上的差值正好等于 0 时， $f_1 = 1$ ，即完全匹配；差值为 ε 时， $f_1 = 0$ ，即匹配度不好；差得比 ε 还大的情况下，就直接记为门限值 0。例如慢跑、 $I=0.3$ 时， $C_{target} = 0.09$ 。模式(4,3)的 $C(M)=1/12\approx 0.083$ ，差距只有 0.007，于是 $f_1 = 1 - 0.007/0.05 = 0.86$ ，是匹配程度比较好的情况。如果强度提高到 $I=0.8$ ， $C_{target} = 0.64$ ，而(4,3)仍然是 0.083，差距是 0.557，已经远超 0.05， f_1 就是 0。即通过这种方式来表征了音乐和运动的适配程度。

②音乐连续性 $f_2(M)$ ：考查音乐在不同段落演进过程中是否具有相对平滑的变化。这是依据 4.3 节的(10)式。如果最大公约数的计算都能符合要求，说明复杂度还是相同的层级，听觉上不会突然变密或变疏，这时 f_2 取较大值；反之就取较小值。这个函数的设定我们先以简化形式，即满足(10)式时设定 $f_2 = 1$ ，不满足时设定 $f_2 = 0.3$ 。仍然用前文的例子，在上一段音乐是(4,3)的情况下，如果下一段还选(4,3)或(4,9)，最大公约数不变，则 $f_2 = 1$ ；如果改成(2,3)，复杂度从 $1/12$ 变化到了 $1/6$ ，那么 f_2 就只有 0.3。

③听觉舒适度 $f_3(M)$ ：考虑的是音程变化过程中是协和还是刺耳。第 3 章中提到，隔 3 个半音是小三度，隔 4 个是大三度，隔 5 个是纯四度，隔 7 个是纯五度，这些都比较协和；而隔 1 个半音是小二度，挤在一起，容易听起来刺耳。因此，通过计算一

段音乐模式里的协和音程占比，来表征听觉舒适度。如果协和音程占比越高， f_3 越接近 1；反之如果小二度这类不协和音程越多， f_3 越接近 0。以 $M(4,3,4)$ 为例，四个音高依次是 0、3、6、9，每一步都正好跳 3 个半音，全都是小三度，所以舒适度较高，这里我们设定 $f_3 = 0.90$ 。

三个目标函数值得到以后，我们就通过动态权重分配的策略，根据运动场景的实时状态来调整各目标的优先级。权重设计基于运动强度 I 和运动状态变化的速率 $v = |I_{current} - I_{prev}|$ 来进行。我们设定：

$$\omega_1 = 0.4 + 0.5 I, \quad \omega_2 = 0.3 + 0.4 v, \quad \omega_3 = 1 - \omega_1 - \omega_2 \quad (15)$$

下面解释(15)式的含义。 ω_1 只跟运动强度 I 有关，当 $I=0$ 时取到最低值 0.4， $I=1$ 时取到最高值 0.9。 ω_2 跟运动状态变化的快慢相关：运动状态几乎不变时取到最低 0.3，最大的变化速率也就是 I 的最大值和最小值之差（为 1）的情况下，此时权重 ω_2 取到最高 0.7。 ω_3 就是 1 减去前两个权重剩下部分。这样的设计把 ω_3 限制在了大约 0.1 到 0.3 之间，这就将舒适度的权重限定在一个不可缺少但又相对跟其他两个目标比较平衡的范围。在运动强度低、变化较慢时， ω_1 、 ω_2 都偏小， ω_3 会自动高一些；而运动强度高、变化较快时， $\omega_1+\omega_2$ 可能已经超过 0.9， ω_3 就会较低。

例如使用 $I=0.3$ 、 $v=0.1$ 进行计算， $\omega_1=0.4+0.5\times0.3=0.55$ ， $\omega_2=0.3+0.4\times0.1=0.34$ ， $\omega_3=1-0.55-0.34=0.11$ 。也就是说，当运动者慢跑且运动强度几乎没变时，所生成的音乐一半以上权重需要考虑的因素为运动强度的匹配，大约三分之一的权重考虑变换的平滑性，剩下的大约十分之一因素为好听不刺耳。

下图给出了动态权重随运动强度 I 和变化率 v 的变化曲线，可以据此并根据实际的运动情况进行调整，以得到最合理的优化目标权重分配。

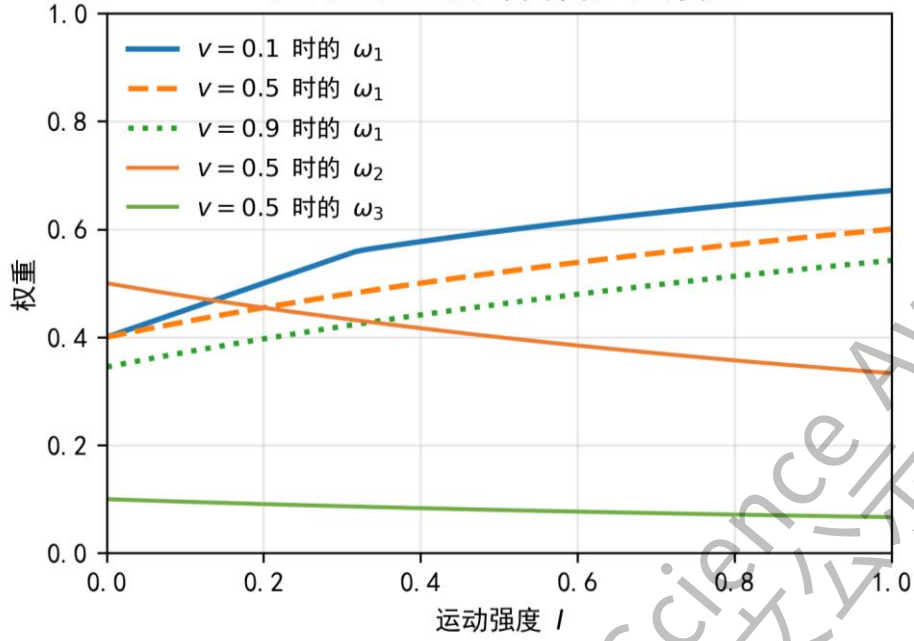


图 6 动态权重随运动强度 I 和变化率 v 的变化。

综上，总的优化目标为：

$$F(M) = \omega_1 f_1(M) + \omega_2 f_2(M) + \omega_3 f_3(M) \quad (16)$$

$F(M)$ 越大，说明音乐模式 M 在当前运动场景下的综合效果越好。当前场景下的最优音乐模式为

$$M^* = \arg \max F(M) \quad (17)$$

还是按照刚才 $I=0.3$ 、 $v=0.1$ 和模式(4,3)的示例数字，计算得到 $f_1 \approx 0.867$ ， $f_2 = 1$ ， $f_3 = 0.90$ ，综合计算得到 $F \approx 0.55 \times 0.867 + 0.34 \times 1 + 0.11 \times 0.90 \approx 0.92$ 。

5 仿真与结果分析

本章对第 4 章的理论模型进行仿真分析。仿真环境为 Windows 操作系统，Python 3 编程环境。基础参数固定为 $N=8$ （八分音符网格）、 $C=12$ （十二平均律）、 $d=4$ 。运动强度 I 取 $[0,1]$ 范围区间。生成元的枚举范围是 $f_r = 2, \dots, 6$ 和 $f_m = 2, \dots, 10$ ，全部组合一共包含 45 组。接下来依次对第 4 章中的内容进行仿真。具体做法是采用手算有代表性的结果，同时用程序仿真把所有可选项都跑一遍，二者相互印证，然后再给出结果分析。

5.1 对于定理 2 子群满足条件的计算分析

在第 4 章定理 2 中已经给出，当且仅当(7)式成立的时候， $M(f_r, f_m, d)$ 是直积群的子群。对 45 组生成元逐个进行检查，判断 $d \cdot f_r$ 是否为 N 的倍数、 $d \cdot f_m$ 是否为 C 的倍数。得出满足要求的有 9 组，占比是 20.0%。

选择若干代表性参数组合进行手动计算。对于(4,3)，可得 $4 \times 4 = 16 \equiv 0 \pmod{8}$ ， $4 \times 3 = 12 \equiv 0 \pmod{12}$ ，两条都成立，所以这个应当是子群。然后再验证一下：根据定义 7 逐点写出(0,0)、(4,3)、(0,6)、(4,9)，走完 4 步后回到(0,0)，可以循环。

再看一个反例(2,7)的情况，可得 $4 \times 2 = 8 \equiv 0 \pmod{8}$ ，节奏维度是符合要求的；但 $4 \times 7 = 28 \equiv 4 \pmod{12}$ ，音高维度不符合要求，因此由定理 2 判定这不是子群。这也跟程序运行的一致。

表 3 列出了所有的 9 组符合要求的生成元。需要说明的是，合格的生成元占比只有 20.0%，说明如果随机选取的话大部分时候不能够得到可循环的类型，因此通过本文的理论来引导生成元选择很有必要。另外，观察表 3 中的内容，可看到例如(4,6)的情况，4 组音位里只有两组不同，片段只能在(0,0)和(4,6)之间来回切换。而这种情况虽然符合要求，但听觉的舒适度就会偏低，因此确实需要后面更多目标和约束的限制来进一步筛选。

表 3 N=8、C=12、d=4 时满足定理 2 的生成元

(f_r, f_m)	$\gcd(f_r, N)$	$\gcd(f_m, C)$	$C(M)$	互异点数	点列
(2,3)	2	3	1/6	4	(0,0),(2,3),(4,6),(6,9)
(2,6)	2	6	1/12	4	(0,0),(2,6),(4,0),(6,6)
(2,9)	2	3	1/6	4	(0,0),(2,9),(4,6),(6,3)
(4,3)	4	3	1/12	4	(0,0),(4,3),(0,6),(4,9)
(4,6)	4	6	1/24	2	(0,0),(4,6),(0,0),(4,6)
(4,9)	4	3	1/12	4	(0,0),(4,9),(0,6),(4,3)
(6,3)	2	3	1/6	4	(0,0),(6,3),(4,6),(2,9)
(6,6)	2	6	1/12	4	(0,0),(6,6),(4,0),(2,6)
(6,9)	2	3	1/6	4	(0,0),(6,9),(4,6),(2,3)

5.2 对于定理 3 复杂度公式和平方映射的可达范围分析

可以使用表 3 来核对定理 3 中(8)式的正确性。以(4,3)为例， $\gcd(4,8)=4$ ， $\gcd(3,12)=3$ ， $C(M)=1/(4 \times 3)=1/12 \approx 0.083$ ，跟表中是一致的。

观察 9 组生成元的复杂度，有三个取值：1/24、1/12、1/6。这是由于最大公约数只取整数，所以得到的复杂度也就只有离散的几个值。

再分析(9)式的 $C_{target} = I^2$ 。I=0.3 时，计算出目标是 0.09，离 $1/12 \approx 0.083$ 很近；I=0.8 时，计算出目标是 0.64。而本组参数中所能得到的最大复杂度只有 $1/6 \approx 0.167$ ，反推计算最大目标时的 I 值 $I^* = \sqrt{C_{max}} \approx 0.41$ ，那么当 $I > I^*$ 时， I^2 就高于任何的合格模式所能达到的复杂度。按照(14)式的处理方法， f_1 对于这些 9 组的值都为 0。图 7 给出了所选模式的 C(M)与 I^2 的差值，其中的竖虚线代表 I^* 。

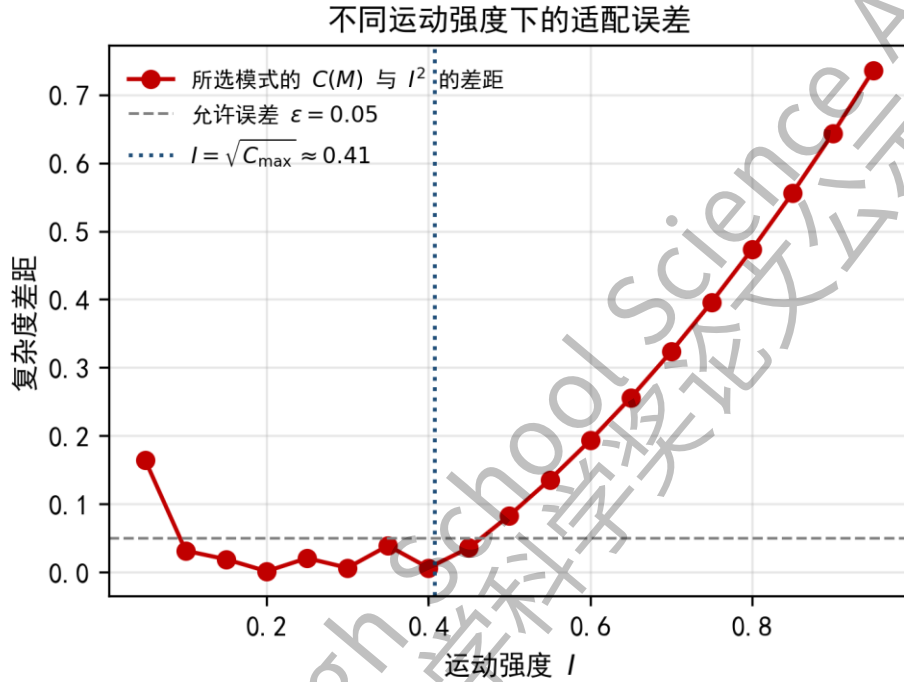


图 7 所选模式的 C(M)与 I^2 的差值

通过仿真分析可以得出，在(9)式的目标设计下，在类似于 N=8、C=12、d=4 这组参数取值的情况下，只有低运动强度（大约 $I \leq 0.41$ ）时，目标计算才能落到实际能取的三种取值里。在运动强度高的情况下， $f_1 = 0$ ，这样优化的目标就变为考虑 f_2 和 f_3 ，这体现了目标优化设计的灵活性。

5.3 平滑过渡变换的计算分析

首先根据(10)式对表 3 里的各项进行评估。复杂度平滑过渡要求两段的两个最大公约数计算分别相等。在表 3 里与(4,3)一致的只有(4,9)，因为二者的最大公约数都是(4,3)，C 都是 1/12。其余的 7 组要么节奏维度上最大公约数从 4 变成 2，要么音高维度上最大公约数从 3 变成 6，C 都会产生非平滑的过渡。

然后根据(13)式进行分析。仍然取上一段为 M(4,3,4)的四个点，缩放平移参数分别取 $k_r = 5$ 、 $k_m = 7$ 、 $b_r = 1$ 、 $b_m = 2$ ，所取的值都跟模数互质。程序的输出为(1,2)、

(5,11)、(1,8)、(5,5)，与第 4.3 节的计算一致。作为对照，如果改成不互质的 $k_r = k_m = 2$ ，计算出来的四个点就变成了(0,0)、(0,6)、(0,0)、(0,6)，这样就只剩下两个不同的点，时间维度上被压缩到了同一拍。图 8 给出了这两种情况的可视化对比。

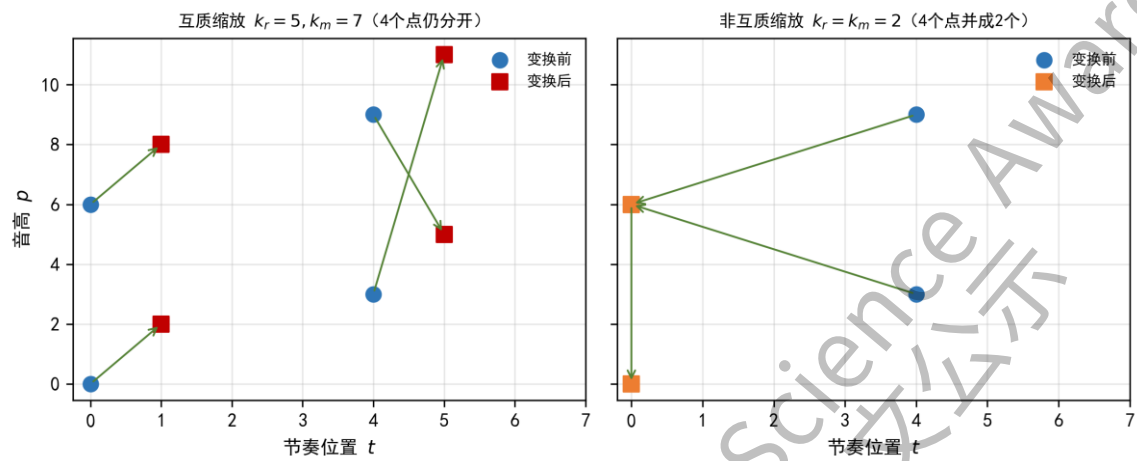


图 8 (13)式的正例与反例

5.4 动态加权选优的分析

在 4.4 节设定的规则里，先用定理 2 筛选出 9 个合格的模式，再按照(14)–(16)式打分，取综合加权评价 F 的最大一组。仿真中 f_2 按 4.4 节的设定取值， f_3 按相邻音程是否协和为计算依据。上一片段固定为(4,3)。

先考虑低强度慢变化 ($I=0.3, v=0.1$) 情况下，把 9 个候选列出，计算其综合分析胡。此时权重为 $\omega_1=0.55$ 、 $\omega_2=0.34$ 、 $\omega_3=0.11$ 。结果如图 9 所示

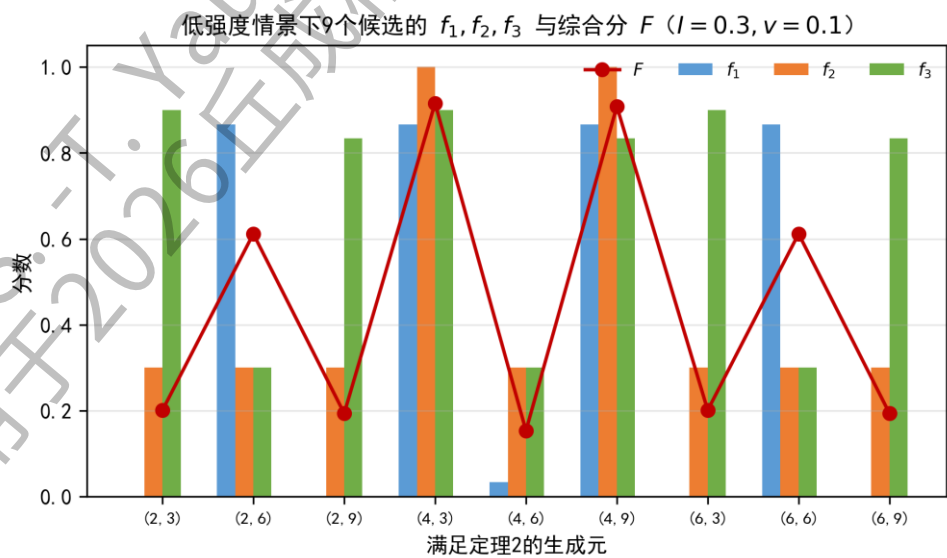


图 9 $I=0.3$ 、 $v=0.1$ 、上一段为(4,3)时，9 个合格模式的三项分数与综合分

从图 9 可以看出，对于 $C=1/6$ 的四组值距离 $I^2 = 0.09$ 超过了 ε ， $f_1 = 0$ ，总分被压到 0.20 附近，排除。而对于 $C=1/12$ 且与上一段同档的(4,3)、(4,9)同时具有 $f_1 \approx 0.867$ 和 $f_2 = 1$ ，这两种情况的总分最高，再进一步比较 f_3 ，其中(4,3)的 $f_3 = 0.90$ 略高于(4,9)的 0.83，即可得到两个的最终目标函数分别为 $F=0.916$ 和 0.908，所选出的跟 4.4 节分析的一致。而对于 $C=1/12$ 的其它两组，即(2,6)、(6,6)， f_2 只有 0.3， f_3 也低（音程大多是三全音），总的 F 只有 0.612。这就表明了 4.4 节设计的模型的有效性，并且能够反映实际应用场景的情况。

接下来进一步分析三种典型情景下的目标函数计算，见表 4。低强度慢变化时，算出的目标 0.09 是在 $1/12$ 附近，所以可以使用非零的 f_1 值计算。中强度中变化 $I=0.6$ 时、和高强度快变化 $I=0.8$ 时，算出的目标 0.36、0.64 都已经高于 C_{max} ，因此 9 组的 f_1 全部都计算为 0；这时(16)式就只包含 f_2 和 f_3 ，计算得出的最优生成元是与上一段同档、并且听觉上相对协和的(4,3)。综合分从 0.92 降到 0.45~0.48，主要是由于 f_1 项的取零带来的影响。

表 4 三种运动场景下的优化结果（上一片段为(4,3)）

情景	I	v	最优生成元	C(M)	I^2	f1	f2	f3	F
低强度慢变化	0.3	0.1	(4,3)	0.083	0.090	0.867	1.000	0.900	0.916
中强度中变化	0.6	0.5	(4,3)	0.083	0.360	0.000	1.000	0.900	0.454
高强度快变化	0.8	0.9	(4,3)	0.083	0.640	0.000	1.000	0.900	0.481

另外，表 4 里高强度情况下算出的 F 反而略高于中强度情况，这是因为(15)式在 v 更大时 ω_2 更大，而最优的一组选择 f_2 项的贡献更多。这体现了实际应用中，运动强度变化越快，越需要适配性强的音乐来实现平滑的过渡。

5.5 连续运动过程及与其它模型的结果对比

在以上分段验证的基础上，我们下面考虑真实的运动场景：热身→慢跑→加速→间歇→冲刺→减速→放松，七段连成一次完整的运动。结果见表 5 和图 10。

表 5 一段连续运动中的选优结果

阶段	I	v	所选生成元	C(M)	I^2	f1	f2	F
热身	0.20	0.00	(4,6)	0.042	0.040	0.967	0.500	0.693
慢跑	0.30	0.10	(4,3)	0.083	0.090	0.867	0.300	0.678
加速	0.45	0.15	(4,3)	0.083	0.203	0.000	1.000	0.415
间歇	0.70	0.25	(4,3)	0.083	0.490	0.000	1.000	0.392
冲刺	0.80	0.10	(4,3)	0.083	0.640	0.000	1.000	0.347
减速	0.40	0.40	(2,3)	0.167	0.160	0.867	0.300	0.645
放松	0.20	0.20	(4,6)	0.042	0.040	0.967	0.300	0.633

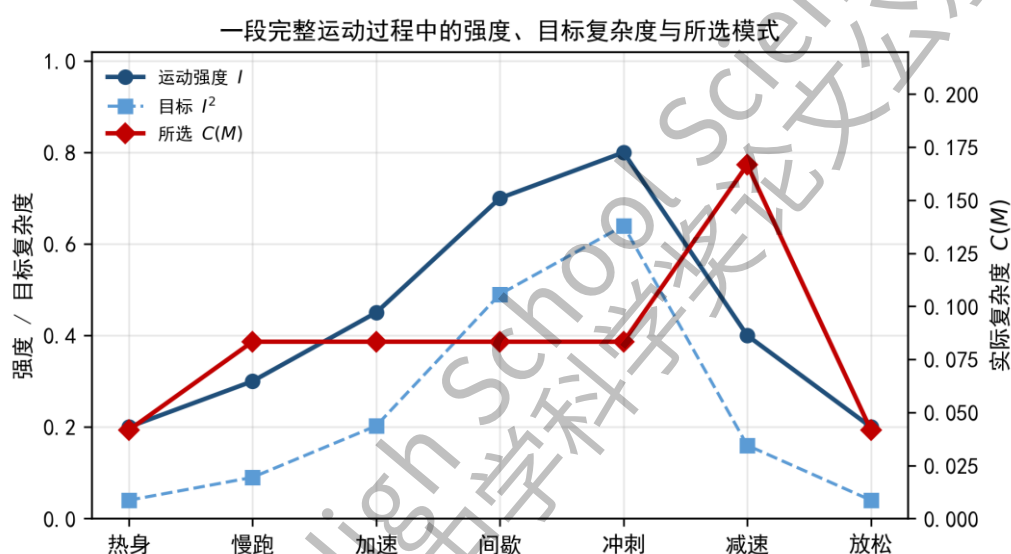
图 10 七段运动中的强度I、目标 I^2 与所选模式的实际复杂度

表 5 体现了在完整运动周期中 4.4 节里的目标权衡的具体过程。热身时还没有上一段， $I^2 = 0.04$ ，最靠近的档位是 1/24，于是选出(4,6)。慢跑时，目标是 0.09，系统就改为选择 1/12 档的(4,3)，尽管跟上一段不同档、而且 f_2 只有 0.3，但这是优化权衡设定的，即适配比连贯更优先。加速以后，I 超过了 I^* ，九组的 f_1 都变成 0，系统不再更换档位，一直留在(4,3)上，冲刺阶段 $f_2 = 1$ 。减速后到 $I=0.40$ 时，目标是 0.16，重新落进 1/6 这一档，系统就改为选择(2,3)： f_1 为 0.867，即使此时 f_2 只有 0.3，总分仍然比继续留在(4,3)要高。放松阶段，又回到了(4,6)。从以上分析可以看出，仿真结果体现了建模过程中(14)–(16)式的合理性：在目标可达的时候关注于运动强度的适配，在目标过高超出门限时，先保证连贯性，运动强度回落后再换回来。

为了进一步说明本文建模的有效性，最后我们把本文结果跟使用别的方法的结果做对照。取 $I=0.3$ 。第一种是采用本文方法，在 9 个合格模式里取 F 最大，取符合条件的子群，计算出 $|C - I^2| \approx 0.007$ 。第二种方法是从 45 组生成元里均匀随机挑，重复 500 次，约 19.6% 的情况下碰巧满足子群条件，平均的 $|C - I^2|$ 约为 0.23。第三种方法是尽可能选择大的生成元，选出 (6,10)，此时 $4 \times 10 = 40 \not\equiv 0 \pmod{12}$ ，不满足定理 2，按 (8) 式计算的复杂度虽能算出 0.25，但点列并不能按小节实现闭合。

图 11 把这三种方法的结果进行比较。首先体现了定理 2 的作用，通过子群性质筛选合格的组合，在此基础上，再进一步用 (16) 式进行比较，适配误差才有机会被压低到 ε 以内。另外，前述结果表明，只有 I 较小、 I^2 还在三档复杂度里面时，误差才会小于 0.05。由结果可知，按照本文的方法可以由最多的符合条件的子群比例，并且计算出的适配误差最小，进一步验证了本文理论模型的有效性。

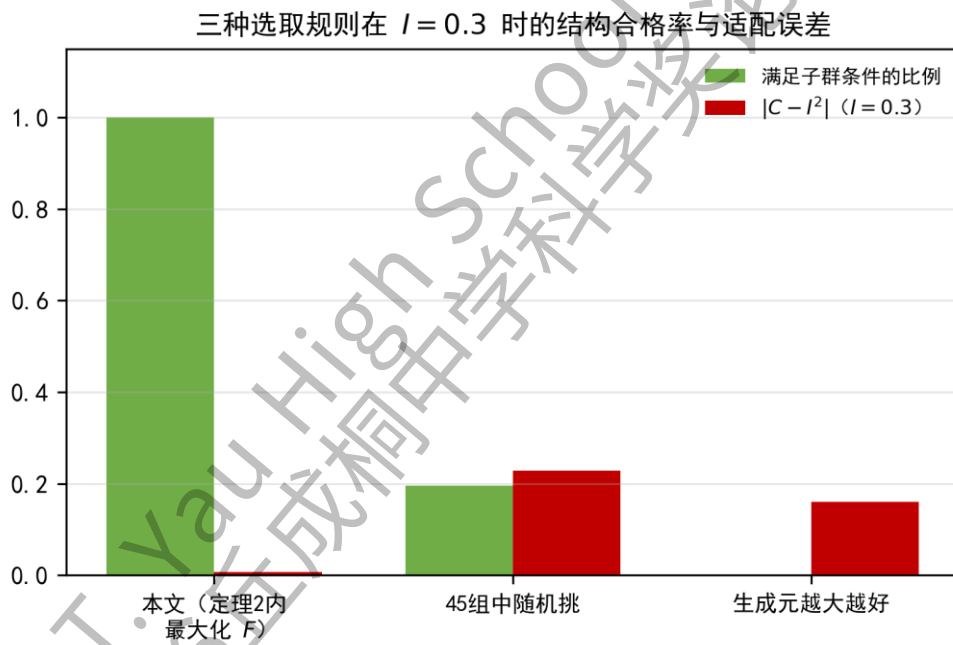


图 11 三种选取规则在 $I=0.3$ 时的子群合格率与适配误差

6 结论

本文针对运动场景下音乐生成的问题，先给出和弦进行的群论简化模型，再把它推广到节奏-旋律二维空间，主要做了以下几方面研究。

第 3 章把大小三和弦写成有序数对，定义新运算，证明集合构成 24 阶阿贝尔群，并用音程变换的样本频率建立一致性函数 E 和多样性函数 H 。对谱例进行统计分析表明，评价指标与直观的和弦进行疏密方向一致。

第 4 章扩展到节奏和旋律两维度，研究对象变为音乐模式 $M(f_r, f_m, d)$ 。定理 2 给出它成为直积群 $Z_N \times Z_C$ 的子群的充分必要条件，定理 3 基于最大公约数定义复杂度，并给出与运动强度之间的映射关系，进一步给出保持复杂度平滑过渡的平移-缩放变换，以及包含运动适配度、音乐连续性和听觉舒适度的动态权重综合评价。

第 5 章仿真分析，在设定的参数下，计算符合条件的生成元比例，评估不同复杂度的特性，验证互质的平移-缩放能够保持音乐模式的性质不变。运动强度不同情况下对优化模型采用灵活调整的方法，以切合实际运动场景情况。将本文方法与其它方法的结果进行比较，进一步验证了本文模型的实用性和有效性。

后续工作可以从以下几方面展开：一是采用更复杂的参数进行仿真，从更多的数据角度来深化对模型跟实际运动场景适配程度；二是进一步深化理论分析，将多目标优化部分使用更复杂的数学模型来建立；三是考虑优化调整模型，扩展应用到其它类型的音乐生成场景。

7 参考文献

- [1] Wright D. Mathematics and Music[M/OL]. St. Louis: Washington University in St. Louis, 2009.
- [2] Amiot E. About the Number of Generators of a Musical Scale[R]. Perpignan: CPGE, 2009.
- [3] 高章涵. 面向钢琴即兴伴奏智能生成的数学建模与数据分析[R]. 第十六届中学生数理化生综合实践活动, 2024.
- [4] 孙维权. 钢琴即兴伴奏入门教程[M]. 上海: 上海音乐出版社, 2010.
- [5] 李大潜. 数学与音乐[M]. 北京: 高等教育出版社, 2015.
- [6] 李大潜. 对称与群[M]. 北京: 高等教育出版社, 2013.