Analyzing Modular Smoothness in Video Game Music

Elizabeth Medina-Gray

KEYWORDS: video game music, modular music, smoothness, probability, *The Legend of Zelda: The Wind Waker*, *Portal 2*

ABSTRACT: This article provides a detailed method for analyzing smoothness in the seams between musical modules in video games. Video game music is comprised of distinct modules that are triggered during gameplay to yield the real-time soundtracks that accompany players' diverse experiences with a given game. Smoothness—a quality in which two convergent musical modules fit well together—as well as the opposite quality, disjunction, are integral products of modularity, and each quality can significantly contribute to a game. This article first highlights some modular music structures that are common in video games, and outlines the importance of smoothness and disjunction in video game music. Drawing from sources in the music theory, music perception, and video game music literature (including both scholarship and practical composition guides), this article then introduces a method for analyzing smoothness at modular seams through a particular focus on the musical aspects of meter, timbre, pitch, volume, and abruptness. The method allows for a comprehensive examination of smoothness in both sequential and simultaneous situations, and it includes a probabilistic approach so that an analyst can treat all of the many possible seams that might arise from a single modular system. Select analytical examples from *The Legend of Zelda: The Wind Waker* and *Portal 2* provide a sampling of the various interpretive gains available through this method.

Received August 2018

Volume 25, Number 3, October 2019
Copyright © 2019 Society for Music Theory

[1] Although video game music has attracted significant scholarly attention in recent years, methods specifically designed for analyzing video game music are as yet relatively sparse.⁽¹⁾ This scarcity of analytical methods is due not only to the newness of the field of video game music studies, but also to the many opaque and slippery aspects of this music. Among other challenging aspects, video game music's source material is typically encoded, and notated scores often do not exist; moreover, because video games are interactive, their sonic output is flexible rather than fixed (Summers 2016a). The issue of interactivity is particularly crucial at this early stage in the field's development, and it raises a pressing analytical question: in short, how do we analyze music that is interactive, flexible, and dependent on a player's individualized actions?⁽²⁾

[2] The analysts who have thus far approached game music's interactivity and flexibility have done so in a variety of ways. Tim Summers, for example, has emphasized the importance of "analytical play" as a critical method of inquiry into a game's music, and has demonstrated a method drawn from diagrammatic film music analysis for representing game music's various interactive cues; Summers' diagrams incorporate descriptions of the music's content and interactive behavior along with interpretation of this music in the context of the game (Summers 2016a). In the course of examining certain video games through the lens of the music game genre, Steven Reale has raised the possibility that some games contain an "intended composition" or "ideal musical

object,” which a player may realize through gameplay; one might, for instance, analyze a game’s musical methods of providing feedback to players regarding their success or failure in realizing such an ideal object (Reale 2014). Elsewhere, I have promoted analytical attention to the separate components—modules—within video game soundtracks, along with the various ways in which these modules may sound together when triggered during interactive gameplay; I have also suggested that qualities of smoothness and disjunction in the seams between musical modules may serve as a productive analytical focus, and I have broadly examined these qualities in the course of analyzing music from select games (Medina-Gray 2014, 2016, 2017). While modularity and smoothness are closely entwined with issues of interactivity and flexibility in video game music, a rigorous and detailed method for analyzing smoothness at the many and varied seams between musical modules is as yet missing from the published literature. This article provides such a method, thereby laying one path by which analysts can approach video game music with its interactive and flexible qualities at the fore.

[3] In this article, I first illustrate some common modular structures in video game music through examples from the game *The Legend of Zelda: The Wind Waker* (hereafter “*Wind Waker*”), and I highlight the importance of smoothness (and disjunction) as an aesthetic and functional aspect of video game music. Next, I outline a detailed method for analyzing smoothness in the seams between sequential and simultaneous modules by focusing on meter, timbre, pitch, volume, and abruptness. The method allows, first, analysis of an individual sequential seam or an individual moment in a simultaneous seam, and second, analysis of all possible seams that might result from a given modular system. (I have found the music21 toolkit to be helpful in conducting this method’s lengthier calculations.⁽³⁾) Analyses of the earlier examples from *Wind Waker* serve as initial illustrations of the method. To close the article, I analyze two additional examples; these final two case studies demonstrate several ways in which close analysis of modular smoothness can deepen and inflect our understanding of games.

An Introduction to Modular Structure and Smoothness

[4] At a basic level, the music in video games is modular: distinct musical modules are stored in a game’s code along with rules for how those modules can be triggered and modified during gameplay. Following the programmed rules and responding to a player’s actions, the computer (that is, game console, personal computer, etc.) assembles the modules in real time—stringing them together and layering them on top of each other—into a soundtrack that accompanies a player’s individualized experience with the game (Medina-Gray 2016). Modularity, in short, allows game music to be dynamic, to change along with a game’s similarly flexible visuals and events, so that just as each player’s playthrough of a game is unique, no two real-time soundtracks for a given game will be the same. In other words, as with any indeterminate music, a single sounding result of a game’s modular music system is just one among other (sometimes many other) possible realizations of the same system.⁽⁴⁾

[5] For instance, **Example 1** maps out a small selection of musical modules and their corresponding triggers in *Wind Waker* (click on a module to bring up its transcribed musical score) that will serve as raw material for the bulk of this article’s analyses. Each of these modules is defined by the music’s behavior during gameplay—for example, if, without any other input, particular music plays and then returns to its beginning and repeats the same material exactly, that music makes up a looping module. Before I provide the larger context for the game and gameplay of which this music is a part (a critical consideration for any eventual multimedia analysis), and before I explain the details of the modular structure that Example 1 represents, consider first how even a small portion of this example raises analytical challenges. **Video Example 1** shows two separate clips of recorded gameplay during which the music shifts from the stacked “Ocean” modules to the “Outset Island” module. (In this and other videos in this article, the recorded gameplay appears on the left-hand side of the window, and annotation illustrating the modules and their triggers as they occur appears on the right.) Although both gameplay clips in Video Example 1 involve the same modules, the timing for the switch between modules depends on an individual player’s actions, and so the precise *seam* between the modules—the precise musical content—in these two clips is quite different. In both clips, “Outset Island” enters with bass instruments that bounce among D $\flat$ and A $\flat$ pitches, but in Clip A this module begins after the “Ocean” modules execute a functional predominant–dominant progression in the key of D major (m. 16 in the “Ocean” score), while in Clip B the “Ocean” modules are in the process of an extended $\flat$ VI (B $\flat$ major) harmony (m. 71) when “Outset Island” enters. Both clips represent equally valid realizations of this modular music system, and both soundtracks are relevant at least for that individual player’s experience during those moments of gameplay. A full analysis of this modular music system should thus equally account for both of

these realizations, along with all other possible realizations of the system. This challenge to incorporate a wide purview in analysis of modular game music forms a core charge for the work I do here. Another challenge lies in the question of what musical content to examine when analyzing multiple modules, and how to examine that content. I highlighted harmonic content in the descriptions above because tonal harmony is a familiar music-theoretical focus, but many other aspects of this music bear consideration as well; whatever analytical focus we choose should be relevant to video game music and the questions we want to ask of it.

[6] With its diverse and detailed—but not overly convoluted—modular music structures, *Wind Waker* provides ample material with which to demonstrate some common modular structures in video game music, as well as to illustrate this article’s main analytical questions and method. *Wind Waker* is a highly acclaimed entry in the long-running *Legend of Zelda* series. This game was developed by Nintendo—with music composed by Kenta Nagata, Hajime Wakai, Toru Minegishi, and Koji Kondo—and was released on the Nintendo GameCube console in 2002.⁽⁵⁾ In *Wind Waker*, as in other *Legend of Zelda* games, players play as the young hero Link on a quest to save the world. **Video Example 2** shows a sample of recorded gameplay from the middle portion of the game, where players are encouraged to sail on a vast ocean, explore various islands, and interact with other characters, objects, and enemies in the course of progressing through the game’s story. Video Example 2 incorporates all of the musical modules mapped in Example 1; this video thus illustrates one possible realization of this modular system. (Clip A from Video Example 1 reappears as a portion of Video Example 2.)

[7] At the beginning of Video Example 2, the player (as Link) is treading water in the ocean, and “Ocean” layer 1 is playing (the video begins during the second half of m. 2 in this module).⁽⁶⁾ “Ocean” layer 1 is a 2’17”-long looping module that plays whenever a player is in the open ocean during the day (when no enemies are nearby), and so serves as the basic score for the ocean environment. The map in Example 1 uses a set of graphical tools that I find useful for representing some common modular structures; looping modules like “Ocean” layer 1 are represented with a cylindrical graphic with an open dot for the module’s starting point (which overlaps with its ending point in such looping cases). Soon after Video Example 2 begins, the player talks to the King of Red Lions—an anthropomorphic boat with the spirit of an ancient king—and the various stages of this conversation trigger brief, pitched modules that sound on top of the continuous score. Non-looping modules like these brief conversation modules are represented with rectangular graphics in Example 1’s map, with both open-dot starting points and closed-dot ending points; a labeled addition sign attached to each starting point indicates that the particular module is added to the soundtrack upon the corresponding trigger, and that the module always begins at its starting point. Next, the player climbs into the boat and sets sail, which triggers the addition of “Ocean” layer 2—another looping module—into the soundtrack. This second “Ocean” layer is synchronized with “Ocean” layer 1, which means that the two modules always align in only one way (see the transcribed “Ocean” score in Example 1); the open dot with an “X” in Example 1 indicates this synchronization. “Ocean” layer 2 can add into (and subtract from) the soundtrack at any point in the course of this module, so the corresponding addition sign (and subtraction sign) in the modular map is attached to this module as a whole rather than to any particular point. (A likely means of executing this structure in the game’s code would have both “Ocean” layers always playing, in synchrony, and layer 2 would be muted or unmuted according to the corresponding in-game triggers.)

[8] As Video Example 2 continues, the player next approaches Outset Island—Link’s home island and one of the many islands that players visit and revisit over the course of the game’s story. This move triggers the switch from the “Ocean” modules to “Outset Island,” a looping module that plays whenever a player is on or near this particular island during the day.⁽⁷⁾ The arrow in the modular map indicates the trigger for this change (“sail into Outset Island area”) and provides some additional details about the switch between modules: The fact that the arrow originates from “Ocean” as a whole (as well as “Ocean Intro”) indicates that the switch can happen at any point during the looping “Ocean” modules (or “Ocean Intro”); the fact that the arrow’s head points to the starting point in “Outset Island” indicates that this new module always begins at its starting point. Additional text next to the arrow specifies that this seam is always sequential (more on this later), that “Ocean” fades out and “Outset Island” fades in, and that there is always at least 0.4 seconds of silence between the modules.⁽⁸⁾ Toward the end of Video Example 2, the player sails back out to the open ocean, triggering a switch from “Outset Island” (which can happen at any point in this loop) to the beginning of the non-looping “Ocean Intro;” layer 2 of “Ocean Intro” is synchronized with layer 1, and both layers sound in this instance because the player is sailing. When the eight-measure-long “Ocean Intro” reaches its end, this ending point triggers the switch to “Ocean” (again, both layers 1 and 2 sound in this instance).

[9] As is the case with all of the examples in this article, I constructed the modular map in Example 1 (and transcribed the corresponding scores) using a “black box” approach; in black box testing, the analyst arrives at a reasonable approximation of the underlying system through repeated examinations of the various products of that system.⁽⁹⁾ Lacking access to the game’s code (which is typical for those outside of a game’s developing team), I relied on repeated gameplay and careful tests of a given portion of the game in order to build a robust representation of the modules and their programmed behavior.⁽¹⁰⁾ Although the results are not necessarily exactly the same as the digital objects actually stored in the game’s memory, they are—critically—highly accurate with respect to what players will hear during gameplay.⁽¹¹⁾

[10] The various triggers in Example 1 illustrate the two typical dimensions of modular structure in video game music: horizontal structure, in which modules switch from one to another (indicated with an arrow in the map), and vertical structure, in which modules layer on top of each other (indicated with addition and subtraction signs).⁽¹²⁾ While the conceptual frameworks of horizontal and vertical structure helpfully reflect modules’ large-scale organization and behavior during gameplay, the analytical focus in this article is on the precise seams created when modules combine—that is, the moments in which the modules’ contents come most directly into contact—and this small-scale focus requires an entirely separate consideration of such seams as either *sequential* or *simultaneous*. In a sequential seam, one module stops completely before the next module begins; in a simultaneous seam, the two modules overlap. Vertical structure thus always produces simultaneous seams, but horizontal structure can produce sequential seams or simultaneous seams, depending on the situation. For instance, all of the trigger arrows in Example 1 produce sequential seams (as noted next to the arrows) because the first module stops entirely before the next module begins. By contrast, in a horizontal situation where one module crossfades into another, the overlap between the two modules would yield a simultaneous seam. With this view, I am suggesting that a crossfade’s simultaneous overlap—rather than its broader horizontal structure—provides the most appropriate frame in which to examine the modules’ precise interrelation, and thus serves as the focus of the current analysis. (A crossfade’s horizontal structure remains relevant, however, and will prompt some special consideration in this article’s analytical methods.)

[11] The distinction between sequential and simultaneous seams is a significant factor in this article’s analysis of modular combinations, especially in consideration of *smoothness* at the seams between modules. Smoothness—as I use the term—is a quality in which two modules fit well together; the modules agree with each other and may even merge into a larger piece of music. The opposite of smoothness is *disjunction*. With disjunction, two modules do not fit well together; they disagree with each other and may remain conceptually separate, despite meeting temporally in the soundtrack.⁽¹³⁾ Smoothness and disjunction here are neutral qualities that describe two modules’ real-time connectedness, similar to integration and segregation of sonic events in auditory streaming (Bregman 1990). A single seam, moreover, can be smooth in some ways but disjunct in others; for instance, the seam between “Ocean” and “Outset Island” in Video Example 2 is smooth in terms of its gradual fade-out and fade-in, but it is also disjunct in its complete change of timbres between the first module and the second.

[12] Musical smoothness and disjunction are recurring concerns for video game scholars and composers alike. On the one hand, smoothness is often lauded by game composers as an aesthetic goal. Richard Jacques, for example, has praised composers for designing music that is “smooth and clean and plays underneath the game in [a] seamless way,” similar to a film soundtrack (Sheffield 2008, 5). Winifred Phillips, in her guide to composing music for video games, has summarized the smoothness aesthetic in this way: “The common wisdom in game audio development includes a belief that it is best to knit together all the music elements of a video game into a seamless experience, without noticeable breaks between the tracks” (Phillips 2014, 53).⁽¹⁴⁾ Game audio scholars Karen Collins and Tim Summers have both posited a connection between musical smoothness and players’ experiences of immersion: Collins has suggested that musical smoothness supports continuity of the game and gameplay, while a “disjointed score” (that is, disjunction) may cause a game to “lose some of its immersive quality” (K. Collins 2008b, 145). Similarly, Summers has suggested that a lack of smoothness in a game’s music may “rupture the integrity of the medium, forcibly reminding the player in an unwanted fashion of the constructed nature of the game,” thus threatening the player’s immersion (Summers 2016b, 79–80).

[13] On the other hand, disjunction between musical modules can support some critical facets of gameplay as well, and some authors balance consideration of disjunction—alongside smoothness—as a potentially positive component of game soundtracks. Phillips, after highlighting the game industry’s emphasis on musical smoothness, encourages composers to consider, for example, whether a pause between music tracks might be

“sometimes better than a smooth transition,” because it might challenge listeners’ expectations of how the music will go, and thus stimulate attention (Phillips 2014, 53). Composer Leonard Paul has pointed out that “a good musical break between music segments” (that is, disjunction during a switch from one module to another) can “clearly indicate to the player that they have entered a new location or game state” (Paul 2013, 65).⁽¹⁵⁾ Disjunction may also support what scholar Kristine Jørgensen calls game audio’s *usability* function, which includes providing feedback for players’ actions, by detaching certain sounds from the rest of the soundtrack (Jørgensen 2009, 158, 160). When musical disjunction makes sense in the context of gameplay—for instance, when it aligns with environmental change or promotes usability—this quality also reasonably supports a player’s immersion in the game (that is, belief in the game world, engagement with the game system, flow, and so on).

[14] Again, disjunction here simply means a sonic separation (or segregation) of modules despite the fact that they occur together in a real-time soundtrack. Experientially, disjunction may startle or surprise players, or draw a player’s attention to the music, but this is not necessarily the case, nor is disjunction necessarily a negative effect (or a negative reflection on the soundtrack’s composition and design). More broadly, issues of experience are messy: players’ expectations of and reactions to game music rely at least partly on their previous experiences with the particular game in question, as well as various inter-textual factors such as genre, time period, and technology.⁽¹⁶⁾ Therefore, although the analytical method in this article draws support from some music perception and cognition literature in order to identify and examine qualities of smoothness and disjunction in modular music, this method does not attempt to model or make claims about universal experiences of smoothness, music, or video games.

[15] This article’s analytical method examines smoothness by comparing material on either side of a modular seam in terms of various musical dimensions, a type of approach that has precedents in both music-theoretical and video game literature. Existing approaches to segmentation analysis—usually of post-tonal pieces—examine various musical dimensions in order to determine how a particular piece of music might be segmented into smaller sections.⁽¹⁷⁾ With video game music, however, “segmentation” is already given through the programmed behavior that defines distinct modules; rather than aiming to illuminate segmentation, analysis of smoothness in video game music instead examines the extent to which the music either enforces or obscures that segmentation. Elsewhere, I have examined certain elements of musical modules in order to analyze dynamic music in particular games (Medina-Gray 2014, 2017); the current article expands and deepens this earlier work. Composer Paul Hoffert also provides a general precedent for the current approach; in his guide for composers of interactive media, Hoffert lays out a system for estimating smoothness across a transition from one module to another by considering the degree of contrast in seven musical elements: volume, tempo, rhythm, key, harmony, texture, and style.⁽¹⁸⁾ Although Hoffert’s system is broadly intuitive and approachable, it is not well suited for a detailed analysis of video game music: the seven elements are not all applicable or clearly defined (for example, “key” presents problems that I examine later, and “style” is a particularly vague concept), the system relies mainly on ad hoc judgments of similarity on a numerical scale, and the system does not provide a means of examining vertically related modules nor multiple possible seams between modules at once. The method proposed in this article, by contrast, provides a comprehensive and detailed means for examining modular smoothness in video game music that can take all of the above situations into account.

A Method for Analyzing Smoothness at Modular Seams

[16] In sequential seams as well as simultaneous seams, smoothness and disjunction are indications of whether and how well two modules fit together. For *sequential* seams, smoothness involves continuity from the first module into the second. Since similarity contributes to continuity within a stream of musical events, smoothness in sequential seams relies at least partly on similarities between the musical content before and immediately after the seam.⁽¹⁹⁾ In particular, to analyze smoothness in sequential seams with listeners’ general perceptual capacities in mind, I here ask how well the material in the one clock-second after the seam (the most immediately “new” material) represents a continuation of the musical material in the five seconds before the seam. This 5+1 seconds framework allows for a consistent view of sequential seams based on time rather than music-specific factors such as number of onsets or measures, and fits into the average perceptual limit of four to eight seconds for auditory short-term memory (Snyder 2009, 107).⁽²⁰⁾

[17] For *simultaneous* seams, smoothness involves agreement during the modules’ overlap. The comparison between materials here is more immediate than in sequential seams, so a consideration of short-term memory

is not necessary. Analysis of simultaneous smoothness thus simply asks how well the two modules agree with each other in the immediate moments at which they are sounding at the same time.

First steps for analysis: Rubrics for individual seams

[18] For both sequential continuity and simultaneous agreement, the current analytical method focuses on five main aspects that regularly yield either smoothness or disjunction in video game music, four of which mainly correspond to familiar dimensions of musical sound: *meter* (or temporal organization of pulses), *timbre* (or instrumentation), *pitch* (which, for these purposes, mainly refers to pitch class content), and *volume* (loudness or softness). The fifth aspect is *abruptness*, which corresponds to the technical treatment of modules as they start and stop (that is, whether the modules enter or exit the soundtrack in a gradual and/or natural manner, or whether they are artificially cut off). These five main aspects are reasonably intuitive—they correspond with my own experience of this music and align with certain suggestions from composers of video game music—and they are based in theoretical and perceptual understandings of how listeners tend to hear sequential continuity and simultaneous agreement in music (all of which is discussed in detail below). **Examples 2 and 3** provide the complete rubrics for analyzing smoothness in sequential and simultaneous seams, respectively; these two rubrics form the basis of the current analytical method. Each row in the rubrics categorizes smoothness and disjunction in a particular aspect (meter, timbre, pitch, volume, and abruptness, with space at the bottom of the rubrics for other aspects as appropriate); strong smoothness falls on the left-hand side of each rubric, strong disjunction on the right, and mild smoothness and disjunction toward the center. The following discussion supports and explains the analytical focus for each aspect, taking each row of both rubrics in turn (first meter, then pitch, and so on). Readers may wish to glance through Examples 2 and 3 now for an overview of the rubrics (and to note their similarities), then keep these rubrics handy while reading the detailed discussion of each aspect below. **Examples 4a, 4b, and 4c** transcribe a few select seams from Video Example 2 (in chronological order), and these examples serve as useful reference points during the following discussion.

[19] *Meter*. Some composers have suggested that temporal aspects like tempo and rhythmic activity can affect smoothness in video game music.⁽²¹⁾ A concept of meter—layers of pulses—provides an especially robust framework with which to approach temporal smoothness in video game music, since meter encompasses and more clearly defines qualities like “tempo,” and much of video game music is metric. David Huron provides a perceptual perspective on meter’s relevance for smoothness (that is, goodness of fit): “research has established that tone onsets coinciding with the most [statistically] common metric positions are judged as ‘better fitting’ with an antecedent metric context” (Huron 2006, 179).⁽²²⁾

[20] As with most of the five musical aspects under consideration here, meter requires a somewhat different consideration in sequential seams (Example 2) and simultaneous seams (Example 3). In sequential seams, the first few onsets in a new module will continue the preceding module best if they continue the prevailing meter. This smoothest metric case occurs, for instance, in Example 4c, where the beginning of “Ocean” continues the eighth note and quarter note pulse streams (among others) established by “Ocean Intro.” (A repeated rhythmic pattern in the percussion—not shown in the transcription—also continues across the seam in this example, providing an additional source of smoothness not covered by meter.) For simultaneous seams, the smoothest metric situation occurs when the onsets in the respective modules support a single meter, as in the two simultaneous “Ocean” layers on the left-hand side of Example 4b.⁽²³⁾

[21] If the meters of the two modules are completely discontinuous, or if they completely disagree, then the metric situation is instead strongly disjunct: In Example 4b, the 140 BPM quarter-note pulse stream in “Ocean” (with 429 milliseconds (ms) between each pulse) does not continue into “Outset Island” (whose quarter-note pulse stream is 119 BPM, with 504 ms between each pulse), nor do any other pulse streams clearly continue, so this seam is as disjunct as possible in terms of meter.⁽²⁴⁾ Gradations in agreement between the meters of two modules can be productively viewed in terms of metric dissonance, after Harald Krebs (1999). In such cases, at least one pulse stream continues across or is supported by two modules, but other pulse streams conflict to yield displacement and/or grouping dissonance.⁽²⁵⁾ I suggest that a conflict involving only one of either grouping or displacement dissonance produces mild smoothness since, in the case of grouping dissonance, some accents (slower pulse streams) still agree, and in the case of displacement dissonance the metric hierarchy is similar (albeit offset).⁽²⁶⁾ A more complex conflict involving both grouping and displacement dissonance more reasonably crosses into disagreement and so produces mild disjunction. As a

general guideline, following Fernando Benadon's analyses of microtiming deviations in early jazz, agreement or alignment among pulses or onsets can be considered to exist if the agreement is to within 50 ms.⁽²⁷⁾

[22] Although much video game music is metric, it is also fairly common for one or both modules in a particular situation to lack meter. The conversation modules in Example 4a, for instance, consist of only one clear attack point each (disregarding the very quick grace notes), but even this single attack point can align or misalign with pulses in the simultaneous module's meter, so an analysis of metric smoothness here focuses on such individual alignment. The strongest metric smoothness occurs when the single-onset module coincides with a strong pulse in the other module's meter, as when the first conversation module in Example 4a coincides with a downbeat of the $\frac{4}{4}$ meter in "Ocean." Mild smoothness arises when the brief module aligns with a weaker pulse (suggesting metric dissonance). A misalignment with metric pulses yields disjunction, as in the second conversation module in Example 4a, which falls between eighth-note pulses in "Ocean." In sequential situations where the first module has no discernible pulse stream(s), the metric aspect simply plays no role in smoothness, since there is no pulse stream that the second module's onsets can continue. In simultaneous situations, even if both modules lack meter, the analyst can still consider whether individual onsets align or not.

[23] *Timbre*. In a brief discussion about ensuring "compatibility" among pieces of video game music, composer Aaron Marks has suggested that "the first key to interactive music is to use the same sound bank and same instrumentation" (Marks 2009, 234). From a perceptual point of view, timbre is a critical factor in auditory scene analysis, whereby some sounds group into auditory streams and segregate from other sounds (Bregman 1990). Timbre is a complex concept and an ongoing area of research, the details of which are beyond the scope of this article.⁽²⁸⁾ A few broad guidelines serve the current analytical purposes, however, starting with one main generalization: if some or all of the instruments in the two modules involved in the sequential or simultaneous seam are the same, then the seam is smooth with respect to timbre; if no instruments are shared, then disjunction results. This consideration of "instruments" includes synthesized sounds, as well as performance techniques (that is, how the sound is produced on an instrument, whether through plucking, bowing, striking, using a mute, etc.).

[24] For sequential seams, the smoothest timbral situation occurs when all of the instruments (and performance techniques) before and after the seam are the same; in other words, the second module continues to use all of the instruments before the seam, and it does not introduce any new instruments. If some (but not all) instruments before and after a seam are the same, as in Example 4c, then the seam is mildly smooth in terms of timbre. For simultaneous seams, the two modules most closely agree in timbre if all of the instruments in one module also exist in the other module; if only some of the instruments (and techniques) in one module are in the other module, then mild smoothness results. For example, if a module with solo piano instrumentation combines simultaneously with another module containing piano and percussion, then the resulting seam will be timbrally very smooth; by contrast, a piano-and-flute module combined with a piano-and-percussion module would be mildly smooth in terms of timbre.

[25] For both sequential and simultaneous seams, entirely different instrumentation or performance techniques across the seam reasonably yield disjunction, and this disjunction can vary in degree. In general, if one or more of the different instruments on either side of the seam are related by means of sound production—for instance, the low brass and trumpets in "Ocean" layers 1 and 2 in Example 4b are both blown brass instruments—then those timbres are typically broadly similar in some aspect, and disjunction can be considered relatively mild. (Timbres with unidentifiable instrumental sources can also be compared by considering components of timbre.) Unrelated instruments, or instruments with different performance techniques, will tend to produce more drastic timbral contrast, resulting in stronger disjunction.⁽²⁹⁾ With this view, for instance, the sequential seam from "Ocean" to "Outset Island" in Example 4b features strong disjunction through timbre; although string basses play in both modules, "Ocean" features a bowed articulation while "Outset Island" uses pizzicato (the other instruments across the two modules are also unrelated, including the percussion not shown in Example 4b). In Example 4a, the simultaneous seams between the conversation modules and "Ocean" layer 1 also feature strong timbral disjunction, since the brief modules' mallet-like synth instrumentation is unrelated to "Ocean" layer 1's strings, brass, and unpitched percussion.

[26] *Pitch*. Several composers have suggested that pitch is an important factor in the combinations of musical modules in video games, and these considerations often cast pitch content in terms of key.⁽³⁰⁾ Such references to key benefit from this historically tested and broadly familiar structural understanding of pitch content, and

tonality or tonal center may indeed reasonably play a role in players' experiences of game music, especially on a larger scale (that is, moving from one key to another). However, a focus on key as the primary descriptor of pitch content raises at least two significant problems when considering modular seams in video game music: First, games often include modules where a designation of key is inappropriate—because the music is tonally ambiguous, fluctuating, or atonal—or imprecise—because the music uses pitches beyond the collection implied by a diatonic key label. For example, *Wind Waker*'s “Ocean” modules (both layers) are overall in the key of D major, but mm. 29–32 and mm. 69–72 include the pitches B $\flat$, F $\sharp$, and C $\sharp$ as part of an emphatic $\flat$ VI– $\flat$ VII–I cadential progression. The two conversation modules in Example 4a don't have “keys” per se, as they consist of only three and one PCs, respectively. A second problem with focusing on key to examine pitch-based smoothness is that other, more detailed information is more directly relevant for determining how well the pitches of two modules fit together in the moments of a sequential or simultaneous seam.

[27] In sequential seams, I propose that the initial pitches of the new module fit especially well with the previous module's prevailing context if these initial pitches repeat pitches that were present in the music before the seam. Huron phrases a similar idea in terms of expectation: “The occurrence of just a single tone moves that tone into the category ‘likely to occur again at some point’” (Huron 2006, 228).⁽³¹⁾ The pitches' specific registers likely play a role in smoothness, and register can certainly be taken into account in analysis, at the analyst's discretion. For the generalized purposes of this article's method, however, it is reasonable to recognize that octave equivalence is often fundamental to pitch organization, and so a repetition of pitch *classes* across the seam also yields very smooth results. Dmitri Tymoczko's concept of *macroharmony* is useful here; a macroharmony is a collection of pitch classes that encompasses all the pitches in a given span of music, and this may be any collection of notes, including but not limited to a diatonic set (Tymoczko 2011, 4). A sequential seam is very smooth in terms of pitch if all of the pitch classes in the one second after the seam are present in the macroharmony in the five seconds before the seam. If none of the pitch classes after the seam are present in the macroharmony before the seam, then the seam is very disjunct in terms of pitch. The remaining situations fall somewhere in between: a seam is reasonably mildly smooth in terms of pitch if most (at least half) of the pitch classes after the seam are present in the previous macroharmony, as is the case in both Example 4b and Example 4c. In Example 4b, out of the D $\flat$ and A $\flat$ after the seam, the D $\flat$ appears (enharmonically, as C $\sharp$) in the macroharmony before the seam;⁽³²⁾ in Example 4c, out of the D, F $\sharp$, and A after the seam, the D and A appear in the macroharmony before the seam. Finally, when relatively few (less than half) of the pitch classes after the seam appear in the previous macroharmony, mild disjunction in terms of pitch reasonably results.⁽³³⁾ This focus on repetition of pitch classes is admittedly reductive, and it leaves out factors related to pitch that might also affect sequential smoothness, such as register or harmonic function (see, for example, the potential expectation of arrival at a D-major tonic harmony that is fulfilled in Example 4c and denied in Example 4b). Overall, however, this article's method gives a broadly applicable means with which to view video game music's widely varying pitch content, and an analyst can always add consideration of other pitch-based components into the analysis, as they become relevant.

[28] In the case of simultaneous seams, the modules' respective pitches directly interact, and some music perception literature supports a basic connection between simultaneous smoothness and consonance. Huron, for example, suggests that “harmonic congruence” (that is, consonance) strongly affects listeners' judgments of how well a concurrent probe tone fits with simultaneous musical material (Huron 2006, 47).⁽³⁴⁾ Similarly, Bregman proposes that consonance helps multiple simultaneous pitches to perceptually fuse into a single chord, and multiple instrumental parts to integrate into a single piece of music (Bregman 1990, 495–96). The pitches of two simultaneous modules reasonably fit together well, then, if their combination creates only consonance; dissonance, by contrast, yields disjunction. Dissonance is here a marker of difference, or separation, between two simultaneous modules' pitches, so the most relevant analytical focus is on the new intervals created by these pitches in combination, rather than the intervals already present in the individual modules by themselves. For instance, in Example 4b, the simultaneous combination of “Ocean” layers 1 and 2 in m. 15 (the third measure in the example) creates no new intervals because all of layer 2's pitches (B and D) also already exist in layer 1; by comparison, layer 2's fleeting C $\sharp$ in m. 13 creates several new intervals in combination with layer 1, including several dissonances (the tritone C $\sharp$ –G, the minor 2nd C $\sharp$ –D, and the major 2nd C $\sharp$ –B).⁽³⁵⁾ In simultaneous seams, smoothness is reasonably reevaluated at each moment of the ongoing combination, so this method treats individual moments during a simultaneous seam separately (although treatment of meter necessarily requires a somewhat wider view).⁽³⁶⁾

[29] Consonance and dissonance are complex qualities that can be affected by a variety of details including tuning, register, upper partials, and context.⁽³⁷⁾ In the interest of generality and approachability, however, this

method reduces out such complicating aspects and broadly categorizes intervals between pitch classes as *consonances* (perfect 4ths and 5ths, and major and minor 3rds and 6ths, in other words, interval classes 3, 4, and 5), *soft dissonances* (major 2nds and minor 7ths, that is, interval class 2), and *hard dissonances* (minor 2nds, major 7ths, and tritones, that is, interval classes 1 and 6).⁽³⁸⁾ For this method, the smoothest situation in terms of pitch occurs when the combination produces no new intervals (assuming equivalence in register, inversion, and enharmonicism) because one module's pitch classes all also appear in the other module (as in m. 15 of "Ocean" layers 1 and 2 in Example 4b) or because one or both modules contain only rests at that point. Moderate smoothness arises when consonant intervals are produced (and no dissonance). Any dissonance produces disjunction: soft dissonances reasonably produce only mild disjunction, while the presence of any hard dissonances yields stronger disjunction. To my ear, hard dissonance is the most significant source of disjunction through pitch. Therefore, given a simultaneous combination that produces both hard dissonance and soft dissonance in the same moment, this method focuses on the hard dissonance only; the more hard dissonance produced in a particular moment, the greater the disjunction. As with the use of macroharmony to gauge pitch-based smoothness in sequential seams, the current focus on intervals in simultaneous seams leaves out other details of pitch that may also affect smoothness; but again, analysts can always decide to treat additional aspects of pitch, as appropriate.⁽³⁹⁾

[30] *Volume*. Hoffert has pointed to similarity in volume at the boundaries of sequential seams as one source of smoothness in video game music (Hoffert 2007, 35–36). More broadly, I suggest that consistent volume around a sequential seam—in the seconds before and after the seam—produces smoothness, while any significant change in volume during that time produces disjunction. For instance, the sequential seam in Example 4c is smooth in terms of volume, while Example 4b, with its fade out to silence and back, is disjunct in this aspect; the significant change in volume in Example 4b supports discontinuity between the music on either side of the seam. (While changing volume produces disjunction in Example 4b, this example's fades and intervening silence also yield smoothness in other aspects, as discussed below.) In simultaneous seams, smoothness results when the volumes of the two modules are the same or similar (that is, the modules occur at about the same dynamics in the mix); this is the case in all of the simultaneous seams in Examples 4a, 4b, and 4c, although dynamics are not specifically labeled in these transcriptions (because the volumes of the modules are all the same). If one module is significantly louder than the other, or if one module increases in volume at the same time as the other module decreases (for example, during a crossfade), then the two modules more clearly disagree in terms of volume, and disjunction results.

[31] *Abruptness*. In horizontally organized systems—where one module switches to another, with either a sequential or simultaneous seam—the abruptness with which the first module stops and the next module begins also plays a role in smoothness. Abruptness arises as an analytical focus perhaps especially because of video games' technological basis; Karen Collins—in the context of a discussion about transitions in video game music—has highlighted a distinction between older practices of hard cuts between modules and more recent tendencies for music to fade and/or crossfade (K. Collins 2008b, 146). Composer Chance Thomas has noted that a long fading sustain at the end of one module can facilitate a "smooth hand-off" to the next module (Thomas 2016, 84). In general, for both sequential and simultaneous seams in horizontally organized systems, if the music proceeds gradually through fades or natural-sounding decays—as in Examples 4b and 4c—then this basic sonic continuity contributes smoothness to the seam. If the music is cut off (as in a hard cut), however, disjunction in this aspect results. Abruptness is less relevant in vertically organized systems than in horizontally organized ones, and so it need not be taken into account in such cases.

[32] *Additional contributions*. The rubrics in Example 2 and Example 3 both leave room for additional sources of smoothness and disjunction that might arise in particular seams—as the analyst deems relevant—starting with but not limited to the suggestions at the bottoms of the charts. For instance, the sequential seam in Example 4b includes 0.4 seconds of silence that buffers the seam (the intervening silence obscures some of the contrast between the music on either side of the seam), providing another source of smoothness.

[33] With the rubrics in Example 2 and Example 3 as a reference, full analysis of individual seams can take advantage of these rubrics' convenient spatial organization. **Examples 5a, 5b, and 5c** analyze particular seams originally transcribed in Examples 4a, 4b, and 4c, respectively. A shorthand grid represents the analytical rubric (for either sequential or simultaneous seams, as appropriate), and cells in the grid correspond to the rubric's categories of smoothness and disjunction in each of the five main aspects; the appropriate category is shaded in for each aspect.⁽⁴⁰⁾ Any additional sources of smoothness or disjunction are indicated with plus signs below the grid.

[34] It is worth noting at this point that this analytical method does not continue on to condense the results for various aspects into a single overall value of smoothness for a seam. Such a compilation would require determining how to weight the various aspects relative to each other, and such weighting reasonably depends on the seam's specific context; while a context-specific approach to weighting can productively serve as an aspect of analysis/interpretation in itself and may be feasible in more linear (determinate) music, in modular video game music this type of analysis would require that an analyst individually determine the weightings for each possible seam.⁽⁴¹⁾ I do not, therefore, wish to weight the various aspects universally for all seams, nor do I wish to try to account for all of the widely varying contexts for seams. Instead, analysis involves examining the various musical aspects for smoothness separately and leaving it at that; a seam as a whole can then be described as "entirely smooth," "predominantly smooth and partially disjunct," "predominantly disjunct and partially smooth," and so on. Such summarizing descriptions allow the analysis to acknowledge sources of smoothness as well as disjunction within a seam, a critical benefit since both qualities may serve important functions in the larger context of the game.

[35] When a player is sailing on the ocean in *Wind Waker*, only one possible seam exists between "Ocean Intro" and "Ocean"—the end of "Ocean Intro" triggers the beginning of "Ocean," and the switch from one set of layered modules to the other cannot happen at any other point during "Ocean Intro"—and so Example 5c represents the full analysis for this system.⁽⁴²⁾ The analysis, in turn, opens space for interpretation: The entirely smooth seam between "Ocean Intro" and "Ocean" effectively makes the connection between these modules appear *seamless*, so that the modules are tied tightly together into one larger piece of music that introduces and represents the adventuresome ocean environment.⁽⁴³⁾ Indeed, the modules are only revealed to be distinct through their behavior (that is, the intro material is not included when the loop repeats). This fusion of the various ocean modules is likely intuitive—even obvious—upon listening, and analysis of the individual musical components of the seam illuminates and nuances this effect.⁽⁴⁴⁾

[36] The analyses in Examples 5a and 5b also reveal potentially illuminating information about these particular seams, and we could at this point interpret ways in which smoothness and disjunction contribute to each of these gameplay situations. Each of these analyses, however, treats only a single seam out of the many seams that can arise from these modular systems. In other words, while valid for a single player's experience of the game, the analyses in Examples 5a and 5b may not be representative of the combinations between these various modules across all possible gameplay scenarios. For a more comprehensive view of modular video game music, the analytical method next takes a probabilistic approach.

Probabilistic Analysis Involving Multiple Seams

[37] The method requires that we analyze—using the smoothness rubrics—*all possible* sequential seams or simultaneous combinations that can arise from the modular system under consideration, and compile these individual analyses into an overall summary. Such a summary conveys the percentages of all possible seams/combinations that contain particular degrees of smoothness in various aspects. Moreover, in systems where all possible seams/combinations are roughly equally likely across all players' experiences, this method's analytical results also represent the likelihood that any given gameplay session will contain smoothness when these modules combine. In short, by taking all possibilities into account, we can contextualize a single seam's qualities as typical, rare, or somewhere in between, and we can view a full range of results from a given modular system at once.

[38] With this wider approach, cells in the shorthand grid are shaded in only if that category applies to every possible seam; if more than one category of smoothness/disjunction is possible for a given aspect, then the grid instead shows the *percentages* of all possible seams that contain those categories. For instance, **Example 6** shows the full analysis for all possible sequential seams while sailing into the Outset Island area. The module "Outset Island" always begins at its beginning, but the sequential seam can occur at any time during "Ocean," or indeed during "Ocean Intro." ("Ocean Intro" and "Ocean" are treated together here because—as the earlier analysis illustrated—the entirely smooth sequential seam between these modules allows them to function together as one larger piece of music. This analysis also incorporates both layers of the ocean's music, since players are most likely to be sailing when the seam occurs.) In this system, the results for meter, timbre, volume, abruptness, and the bonus to smoothness through the buffering silence are all consistent regardless of when during "Ocean Intro" or "Ocean" the seam occurs, and so these portions of the analysis are the same as the analysis of the single seam in Example 5b. (In other words, "Outset Island" can never continue the meter in "Ocean" or "Ocean Intro," the instruments on either side of the seam are always very different, and so on.)

The results for pitch vary, however, and the full analysis (for which a detailed method appears below) enumerates this variety: 13% of the possible seams contain strong smoothness through pitch (that is, both pitch classes after the seam appear in the macroharmony before the seam), 57% contain mild smoothness (that is, at least half (one) of the pitch classes after the seam appears in the macroharmony before the seam), and 30% of the seams contain strong disjunction (through a complete lack of pitch-class continuity).

[39] To arrive at a complete probabilistic analysis of sequential seams as in Example 6, an analyst would hypothetically need to examine every aspect in every seam individually. In practice, however—and as is the case in Example 6—such close examination is often only necessary for the pitch domain. For many modules, aspects like meter, timbre, and volume are consistent for long spans of the module, meaning that smoothness contributions from these aspects are typically the same regardless of when in that span a seam occurs; similarly, abruptness usually applies consistently to all possible seams from a single trigger. Pitch content in most modules, however, fluctuates with time, which necessitates an examination of pitch in every possible seam.

Example 7 provides detailed steps for conducting a probabilistic smoothness analysis for sequential seams in the domain of pitch only (although this method could be modified if needed to apply to other musical aspects). When dealing with large modules, the steps in Example 7 can become too unwieldy and time-consuming to complete by hand, so I developed computer programs using the music21 toolkit and the programming language Python to automate the process. For the modular system analyzed in Example 6, the earliest possible point at which a seam can happen is approximately three seconds (seven quarter notes) after the beginning of “Ocean Intro” (which I determined through repeated testing during gameplay). Seams in this system are analyzed at every subsequent 16th note as a reasonably fine approximation of the infinitely close number of possible time points. In order to treat every possible seam with unique preceding material (in the 5-second window before the seam), points in the first five seconds of “Ocean” are treated twice: once when “Ocean” first begins (and the preceding context therefore includes some of “Ocean Intro”) and again when the modules have looped (and the same point’s preceding context now includes the looping modules’ end).

[40] The analytical results in Example 6 reveal that every possible seam in the move from ocean to Outset Island involves smoothness as well as disjunction, including at least mild smoothness through pitch in the majority of the seams. With respect to pitch, in other words, although the large-scale change in key from D major ocean music to the D $\flat$ major “Outset Island” is a fairly distant tonal move, the very limited pitch class content at the beginning of “Outset Island” (and especially the D $\flat$ enharmonic with C $\sharp$) often provides at least a partial connection to the music immediately before this change. In this ocean/island travel, both smoothness and disjunction serve important functions for the game. Smoothness here maintains some continuity across the soundtrack, fulfilling a basic aesthetic for musical smoothness and mirroring the continuous gameplay and visuals at these moments in the game. At the same time, disjunction distinguishes between the two neighboring pieces of music and therefore their associated environments. Indeed, this musical disjunction is the only indication that a player has crossed the boundary between ocean and island, since the game provides no other cues (visual, aural, or tactile) to mark this boundary. In short, the musical disjunction is entirely responsible for defining the game’s virtual landscape in this case.⁽⁴⁵⁾

[41] For simultaneous seams, probability calculations come into play in two types of situations: analysis of a single simultaneous seam, and analysis of all possible simultaneous seams in a given system. First, since the simultaneous smoothness rubric (in Example 3) treats only a single moment at a time, and since simultaneous seams typically last longer than a single moment, complete analysis of a single simultaneous seam will necessarily take all the moments of the seam into account. Here, the analysis conveys the *proportions* of the seam that contain various degrees of smoothness and disjunction. This single-seam analysis is especially useful when considering vertically organized modules that can only combine in one way, such as the synchronized “Ocean” layers in *Wind Waker*. **Example 8** shows the analysis of the single seam between the two “Ocean” layers. Here, smoothness through meter and volume are consistent throughout the modules’ lengths, but since timbre and pitch content vary, some parts of the simultaneous seam are smooth through pitch and/or timbre, while other parts are disjunct; the numbers in these portions of the analysis represent the percentages of the seam’s entire duration that contains the corresponding degrees of smoothness or disjunction. Regarding timbre: the trumpet (mm. 1–32) and horn (mm. 35–40) in layer 2 are different but related to the simultaneous low brass in layer 1, yielding mild timbral disjunction; when the violins take over the melody in layer 2 (mm. 41–72), however, these (synthesized) instruments sound to my ear exactly the same as the strings in layer 1, so this portion of the seam (along with the ten measures of rests in layer 2, making up 52% of the seam in all) produces strong timbral smoothness. Regarding pitch: the vast majority of the seam’s length features

smoothness, including 84% of the seam in which no new intervals are created (either because all of layer 2's pitches at that moment also already exist in layer 1, or because layer 2 contains only rests at that point). Disjunction through pitch (that is, dissonance) occurs, but it is relatively sparse and mild. Again, computational programs using music21 were useful in streamlining these calculations for pitch.

[42] The comprehensive view in Example 8 clarifies the degree to which the two "Ocean" layers fit together across their lengths. Smoothness through meter, volume, often pitch, and sometimes timbre ensures that the two layers work together simultaneously as a single piece of music. Indeed, in the entirely smooth portions of the seam, the two layers merge especially closely, so that I sometimes find it difficult to distinguish between the two layers when both are playing. When it is present, disjunction through timbre and occasionally pitch maintains some distinction between the layers (as between instruments in an ensemble), so that the presence and absence of layer 2 in the ongoing soundtrack can reflect a player's interactions with the virtual world.⁽⁴⁶⁾ Sailing—which features both accompanying layers—is thus musically richer, more melodic, and more rhythmically interesting than not sailing. At the same time, through predominant smoothness in the simultaneous seam that ties the two layers together, the soundtrack clarifies that sailing is part of (or a variation on) the overall ocean environment.⁽⁴⁷⁾ Such predominant smoothness among synchronized layers is reasonably common in games with layered soundtracks.⁽⁴⁸⁾

[43] For systems where more than one simultaneous seam is possible between a set of modules, the analysis more accurately shows the *likelihood* that the various aspects will produce smoothness or disjunction at a given moment in the multiple possible seams. For instance, **Example 9** shows the full analyses for the first two brief modules during the conversation with the King of Red Lions on the ocean (for which Video Example 2 provides one instance); these brief modules can occur over any point in the first layer of "Ocean Intro" or "Ocean" (the second layer cannot sound during this point because a conversation cannot happen while sailing) and the analyses take all of these possibilities into account.⁽⁴⁹⁾ Regardless of when either of the brief conversation modules occurs, each moment of its combination with the ocean's music is always very disjunct in terms of timbre and smooth in terms of volume; the results for meter and pitch vary, however, depending on when the conversation modules occur. First, the single onsets for these brief modules *can* agree with the simultaneous module's meter.⁽⁵⁰⁾ In 12% of the possible places where the conversation modules can occur during "Ocean Intro" and "Ocean," the brief module's onset aligns with a strong beat (the first or third beats in the measure, ± 50 ms); 35% of the time the onset aligns with a weaker quarter or eighth-note pulse (± 50 ms), and the remaining 53% of the time the onset disagrees entirely with the simultaneous module's meter. **Example 10** provides the steps for calculating metric smoothness in situations like Example 9, where two modules can agree metrically (because one of the modules has only one onset, or because the two modules contain pulse streams whose inter-onset intervals differ by less than 50 ms) and the triggered module can occur at any time; these steps apply equally for sequential and simultaneous seams.

[44] **Example 11** provides the steps for analyzing smoothness in pitch for all possible moments across all possible simultaneous seams. Since the analytical method focuses on individual moments in simultaneous seams, this probabilistic view is able to treat momentary combinations in isolation and then compile the results with respect to the durations of the sonorities involved. It is not necessary, for example, to fully analyze the proportions of smoothness and disjunction across every possible simultaneous seam and then compile those results. Once again, programs using music21 were helpful for automating these calculations. The analyses in Example 9 reveal this modular system's possible results through pitch: The first conversation module almost always yields disjunction in combination with the ocean's music—and often (46% of the time) the disjunction is moderate or strong, with two or more hard dissonances; it is also possible—although extremely rare—for this brief module to produce strong smoothness through pitch (see, especially, m. 8 of "Ocean Intro," one of the very few spots in the ocean's music where this conversation module's pitches C, D, and G all already exist). The second conversation module—with its lone G—much more readily (34% of the time) produces smoothness with the ocean's pitches, and while disjunction is overall more likely, such disjunction is relatively mild.

[45] Overall, the conversation modules' reliable disjunction with the ocean's concurrent music—always through timbre, and frequently through meter and pitch—keeps these modules separate from the ongoing musical material. Although they are built from pitched tones, the predominantly disjunct context casts these brief modules more clearly as sound effects, distinct from the ocean's music and able to provide clear feedback by aurally marking a player's progression through interactable dialogue. At the same time, some smoothness allows these sounds to at least coexist with the ocean's music in the game's overall soundtrack. Moreover, in

the uncommon but possible instances where a combination is predominantly smooth, a conversation module reasonably veers from sound effect into music, enriching the musical soundtrack and tying a player's dialogue-centered actions at least momentarily into the musical environment of the game.

[46] The analytical method established here allows analysts to examine an extensive variety of modular music systems in video games. Thus far, I have treated particular examples from *Wind Waker* as detailed illustrations of the method in order to demonstrate the relevance of this analytical approach in relation to gameplay. Yet the diversity of modular music systems and gameplay situations in video games is vast—*Wind Waker*, for instance, contains many interesting musical systems beyond those treated above—and ample opportunities for analytical exploration exist. With the goal of exploring some of those opportunities now, the remainder of this article analyzes two examples that raise additional considerations relating to musical structure, gameplay, and, in the second example, a game's published soundtrack. The first example expands the ocean environment from *Wind Waker* (already treated above) into new ludic territory.⁽⁵¹⁾ The second example leaves *Wind Waker* to examine a situation in the game *Portal 2*.

Analysis of Ocean Combat in *The Legend of Zelda: The Wind Waker*

[47] Out on the open ocean in *Wind Waker*, enemies sometimes appear, and the ocean environment becomes dangerous. Minor enemies are especially common, and while these enemies are relatively weak (in that they are capable of dealing only small amounts of damage to Link, and players can defeat them with only a few blows), these enemies nevertheless pose significant threats to Link and therefore to players' progress and success in the game. The enemies try to attack Link, and if he takes too much damage, he dies, and players are required to go back to an earlier point in the game in order to keep playing. **Example 12** shows a map of the musical modules that accompany encounters with minor enemies on the ocean, along with smoothness analysis for the two horizontal shifts in this system: first from "Ocean Intro" or "Ocean" to "Early Combat" (with a crossfade and thus a simultaneous seam), and then from "Early Combat" to "Late Combat" (with a sequential seam). **Video Example 3** shows an instance of gameplay involving these modules, with accompanying annotation. Example 12 and Video Example 3 do not show or analyze the subsequent switches in this system, from "Early Combat" or "Late Combat" back to "Ocean."⁽⁵²⁾

[48] The first horizontal shift happens when a player enters the area near an enemy (for example, a shark-like monster, in the case of Video Example 3), triggering a crossfade from the ocean's music into the beginning of "Early Combat." If Link is in the boat when the switch occurs, then both layers of "Early Combat" play, and if not, then only the first layer of "Early Combat" sounds; the analysis of this simultaneous crossfade in Example 12 takes all possibilities of modules and timing into account.⁽⁵³⁾ This situation also illustrates the importance of examining musical modules as they actually sound during gameplay: to my ear, the very brief overlap during the beginning of "Early Combat" most clearly involves this music's snare and synth voice parts up through about the first half of m. 1, but the bass in "Early Combat" is not actually audible until just after the seam; so while the bass timbre does play a role in this combination—essentially through a sequential seam—the analysis of simultaneous smoothness through pitch treats only the synth voice's initial pitches (C, E $\flat$, and F).

[49] Although this horizontal shift into combat can produce many different simultaneous seams, the resulting degrees of smoothness and disjunction are very consistent: smoothness through timbre and abruptness at this seam always supports basic continuity across the soundtrack, while strong disjunction in meter and pitch along with disjunction in volume (because of the crossfade) supports the change in environment. I have marked an additional source of disjunction because "Ocean Intro" and "Ocean" are predominantly consonant, while "Early Combat" is heavily dissonant, yielding another significant contrast across the seam.

[50] *Wind Waker* provides only subtle visual cues to reflect the change in ocean environment from safety to danger: the camera pulls back slightly, and the enemy may not appear immediately—if at all—in the field of vision. Yet this shift in environment requires players to switch to a different set of behaviors (from free exploration/travel to fight-or-flight) or risk ludic setbacks. In the absence of strong visual cues, reliable and strong musical disjunction thus plays an especially useful role in indicating this change in environment.⁽⁵⁴⁾

[51] From the looping, four-measure-long "Early Combat," the next horizontal shift triggers when the first combat blow occurs (when the enemy hits Link, or when Link hits the enemy). Upon this trigger, the music waits until the end of m. 2 or m. 4 in "Early Combat" and then executes the shift to "Late Combat." Because there are only two points in "Early Combat" where the seam can happen (and few variations in the layers that

can be involved), the number of possible sequential seams that can arise from this shift is very limited. In all cases, the seam is smooth in every aspect, including strong metric smoothness—even hypermeter is guaranteed to be continuous—and additional smoothness through continuing rhythmic patterns.⁽⁵⁵⁾

[52] “Early Combat” and “Late Combat” thus fuse into a single larger piece of music that flexibly progresses with specific gameplay events in this dangerous environment. “Early Combat” is ominous but relatively understated, with few instruments and highly constrained (repeated) harmonic and rhythmic material. “Late Combat” initially continues the bass, synth voice, and snare parts from “Early Combat,” but this later loop is over a minute long, and it incorporates more instruments, as well as increasingly intense melodic and harmonic content. The resulting musical system is able to account for a variety of gameplay experiences: For example, players might flee from enemies before taking or dealing any damage, essentially skimming through only the early and relatively mild stage of a combat encounter before returning to the safe ocean environment and its music. Players might fight an enemy and defeat it quickly, in which case the music progresses to “Late Combat” and then returns to the safe ocean music before much of this longer combat music loop has had a chance to play. Or, even though these enemies are relatively weak, players might repeatedly fail to avoid an enemy’s attacks or to defeat an enemy, leading to a lengthy combat encounter, where the later portions of “Late Combat” might reinforce players’ fears for Link’s health as well as frustration with their own gameplay performance.

[53] *Wind Waker*’s music can thus contribute both clarity and nuance during gameplay in the ocean environment. The several reliable sources of disjunction between ocean and combat music emphasize the switch from safety to danger, while the thorough smoothness in the switch from early to late combat yields a larger piece of music with especially dynamic properties. As players become embroiled in combat, the music can reflect (and encourage) an increase in the encounter’s intensity.

Analysis of Repulsion Gel Bouncing in *Portal 2*

[54] *Portal 2*, the highly rated sequel to 2007’s *Portal*, was developed by Valve Corporation and released in 2011 for personal computers, PlayStation 3, and Xbox 360, with music composed by Mike Morasky. In *Portal 2*, within the test chambers of a decaying subterranean laboratory, players must interact with various physics-based mechanics—including the eponymous portal gun—to complete increasingly difficult spatial puzzles. One of these mechanics is a blue gel that, when spread onto a floor or wall and jumped upon, launches the player-controlled character, Chell, into the air, with subtle musical material accompanying each bounce. Players first encounter this blue repulsion gel in a test chamber bearing the label “1953-01” on the exterior wall. The consistent score for this area is a sparse loop of synthesized bass and treble lines that arpeggiate mostly dominant-to-tonic pairs of harmonies in an overall chromatic (tonally wandering) progression (click on “1953-01” in **Example 13**’s map to see a transcription of this module alone).

[55] Reale has already analyzed some of *Portal 2*’s music in relation to the game’s story and design. For instance, natural and artificial elements in the game’s title music may reflect the duality of these same elements in the character of GLaDOS, a wryly sadistic AI who serves as the series’ main antagonist. Some of the game’s later music—including the music from room 1953-01 and subsequent test chambers, titled “You Are Not Part of the Control Group” on the game’s published soundtrack—builds upon and varies earlier music, much as later puzzles build upon earlier gameplay concepts (Reale 2016).⁽⁵⁶⁾ This existing work on the music of *Portal 2* does not, however, treat the especially dynamic aspects of this music during gameplay. By examining the modular music system in room 1953-01 in more detail, we can illuminate even further depths in the relationship between music and game.

[56] Example 13 shows a map of the musical modules that accompany gameplay in room 1953-01, together with smoothness analysis for all possible simultaneous combinations between synchronized layers.⁽⁵⁷⁾ **Video Example 4** provides an instance of gameplay in this chamber, with annotation of the modular system.⁽⁵⁸⁾ When a player bounces against a blue-gel-covered surface, one of three synchronized “Bounce” modules (apparently randomly selected) adds to the continuous “1953-01,” with an increase in volume that mirrors Chell’s ascent; on her descent, the added layer’s volume decreases, and the layer is removed from the soundtrack at the end of the bounce. The three “Bounce” layers are all slight variations of the same musical material: each layer contains the same flute-like synthesized timbre and the same fluttering 32nd notes (the same rhythm and overall contour) that arpeggiate harmonies from the corresponding measures in “1953-01.” Only slight differences distinguish the three modules: “Bounce” layer 2 is in a higher register than the other two modules, and “Bounce” layer 3 uses different pitches at either the highest or lowest points of each

measure's arpeggios. The three "Bounce" layers thus readily occupy a single larger category, allowing subsequent bounces to group together, while also providing some sonic variety for each bounce.

[57] Several aspects yield smoothness during the simultaneous combination of each possible "Bounce" layer with "1953-01": the "Bounce" timbre sounds to my ear to be the same as the timbre for the eighth-note treble line in "1953-01," the simultaneous pitches do not produce dissonance the vast majority of the time, and the volume of two simultaneous layers is the same at least around the peaks of each bounce (roughly 50% of the time). Such reliable smoothness means that the sonic accompaniment for bouncing reasonably fits into the musical space of the test chamber. But this system of synchronized layers also features significant sources of disjunction (more so than *Wind Waker*'s "Ocean" layers analyzed in Example 8, for instance): the dynamic, physics-dependent rules for bouncing mean that volume for the two layers is significantly different roughly 50% of the time, and the resulting divergent sonic behavior (the fact that the "Bounce" layers get louder and softer, sometimes several times in a row, and independent of the consistent "1953-01") adds a further source of disjunction. Metrically, grouping dissonance (3:2) as well as displacement dissonance (by one 32nd note) keeps the "Bounce" layers' quick arpeggios in a constant state of not-quite-alignment (mild disjunction) with "1953-01." Furthering Reale's characterization of *Portal 2*'s music—and GLaDOS—as an algorithmic process "run amok" (Reale 2016, pt. 1, 3:10), the highly patterned and logical yet off-kilter metric relationship between "Bounce" layers and "1953-01" could seem to arise from the demented persona of the game's antagonist. At the very least, this particular mix of musical agreement and disagreement is a fitting accompaniment to gameplay in an unsettling environment in which science and physics are warped for challenging and nefarious ends.

[58] The "Bounce" layers' close connection to the act of bouncing during gameplay, moreover, opens further interpretive space that reveals interesting tensions with *Portal 2*'s published soundtrack. In the track "You Are Not Part of the Control Group," the module "1953-01" plays while the "Bounce" layers enter and exit the mix very gradually—without adhering to the volume profiles that would result from bouncing in the game—and multiple "Bounce" layers often sound at once (an impossibility in the game's modular system). In other words, this track treats the modules as freely equivalent stems in a mix, that is, as material for arrangement into a larger (and fixed) aesthetic whole. During gameplay, by contrast, the rules that govern the "Bounce" modules' appearance produce additional simultaneous disjunction and keep the "Bounce" layers reasonably distinct from "1953-01," such that I hear these layers not as fully integrated components of a larger piece of dynamic music, but rather, in this gameplay context, as something more akin to sound effects. These fluttering arpeggios that rise and fall in volume might then serve as (synthesized, filtered) approximations of air rushing against ears, or else of a physiological reaction to the bouncing activity (for example, Chell's surging pulse and/or adrenaline).⁽⁵⁹⁾ Since this bouncing sound also produces significant smoothness in simultaneous combination with the ongoing "1953-01," however, this sound is reaffirmed as musical, and is situated within the virtual environment in a way that, say, a more realistic sound of rushing air would not be. The modularity in *Portal 2*'s soundtrack thus contributes—but only during gameplay—to a particularly musical physicality, a significant effect in a game that hinges on manipulations of and motions through a virtual space.

Conclusion

[59] Modular smoothness is a critical facet of video game music, and the method outlined in this article provides terminology and tools with which to examine this facet in fine detail. The probabilistic approach is especially helpful, given that video game soundtracks can play out in a multitude of ways; an analysis should be able to account for all possible results of a particular modular system, and this method provides one means of doing so. Although tailored to video game music (for example, in its consideration of the technical aspect of abruptness), this approach may also productively inform analyses of other modular music outside of video games, especially music where smoothness is of analytical interest.

[60] The examples analyzed above highlight a variety of potential effects resulting from modular smoothness and disjunction in particular gameplay contexts. Smoothness might fulfill a basic aesthetic ideal, suggest continuity across gameplay, tie modules tightly together into a single dynamic piece of music, or suggest that disparate elements (for example, sound effects) at least belong in a game's musical environment. Disjunction might promote (or even define) a change in gameplay environment, or clarify the status of particular sounds as distinct from the musical score; and the effects of smoothness and disjunction are by no means limited to those identified here. My analyses also point to the wealth of interpretive gains that become available when we recognize smoothness and disjunction, account for their likelihood, and then consider their effects together

with all the rich specifics of a game: its story, visuals, physicality, and environments, as well as the behaviors it encourages and affords in its players. Although the examples selected for analysis are not representative of all video game music, they are indicators of the expansive variety of musical situations in games, a vast analytical landscape awaiting exploration. Overall, then, this article serves as one signpost toward a fuller understanding of video game music, and provides one robust set of tools with which to approach this uniquely challenging material.

Elizabeth Medina-Gray
Ithaca College
School of Music
953 Danby Rd.
Ithaca, NY 14850
emedinagray@ithaca.edu

Works Cited

- Benadon, Fernando. 2009. "Time Warps in Early Jazz." *Music Theory Spectrum* 31 (1): 1–25.
- Berndt, Axel. 2009. "Musical Nonlinearity in Interactive Narrative Environments." In *Proceedings of the International Computer Music Conference (ICMC)*, 355–58. MPublishing, University of Michigan Library.
- Bregman, Albert S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. The MIT Press.
- Cheng, William. 2014. *Sound Play: Video Games and the Musical Imagination*. Oxford University Press.
- Childs, IV, G. W. 2007. *Creating Music and Sound for Games*. Thomson Course Technology PTR.
- Collins, Karen, ed. 2008a. *From Pac-Man to Pop Music*. Ashgate.
- . 2008b. *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design*. The MIT Press.
- . 2013. *Playing with Sound: A Theory of Interacting with Sound and Music in Video Games*. The MIT Press.
- Collins, Nick. 2008. "The Analysis of Generative Music Programs." *Organised Sound* 13 (3): 237–48.
- Duane, Ben. 2012. "Texture in Eighteenth- and Early Nineteenth-Century String-Quartet Expositions." Ph.D. diss., Northwestern University.
- Gibbons, William. 2017. "Music, Genre, and Nationality in the Postmillennial Fantasy Role-Playing Game." In *The Routledge Companion to Screen Music and Sound*, 412–27. Routledge.
- Hanninen, Dora A. 2012. *A Theory of Music Analysis: On Segmentation and Associative Organization*. University of Rochester Press.
- Hasty, Christopher. 1981. "Segmentation and Process in Post-Tonal Music." *Music Theory Spectrum* 3: 54–73.
- Hoffert, Paul. 2007. *Music for New Media: Composing for Videogames, Web Sites, Presentations, and Other Interactive Media*. Edited by Jonathan Feist. Berklee Press.
- Huron, David. 1991. "Tonal Consonance versus Tonal Fusion in Polyphonic Sonorities." *Music Perception* 9 (2): 135–54.
- . 1994. "Interval-Class Content in Equally Tempered Pitch-Class Sets: Common Scales Exhibit Optimum Tonal Consonance." *Music Perception* 11 (3): 289–305.
- . 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. The MIT Press.
- Jørgensen, Kristine. 2009. *A Comprehensive Study of Sound in Computer Games: How Audio Affects Player Action*. Edwin Mellen Press.

- Kameoka, Akio, and Mamoru Kuriyagawa. 1969. "Consonance Theory Part I: Consonance of Dyads." *The Journal of the Acoustical Society of America* 45 (6): 1451–59.
- Krebs, Harald. 1999. *Fantasy Pieces: Metrical Dissonance in the Music of Robert Schumann*. Oxford University Press.
- Křenek, Ernst. 1940. *Studies in Counterpoint Based on the Twelve-Tone Technique*. Schirmer.
- Krumhansl, Carol L., Gregory J. Sandell, and Desmond C. Sergeant. 1987. "The Perception of Tone Hierarchies and Mirror Forms in Twelve-Tone Serial Music." *Music Perception* 5 (1): 31–77.
- Krumhansl, Carol L., and Mark A. Schmuckler. 1986. "The Petroushka Chord: A Perceptual Investigation." *Music Perception* 4 (2): 153–84.
- Lehman, Frank. 2017. "Methods and Challenges of Analyzing Screen Media." In *The Routledge Companion to Screen Music and Sound*, edited by Miguel Mera, Ronald Sadoff, and Ben Winters, 497–516. Routledge.
- Lynne, Bjorn. 2004. "A DirectMusic Case Study for Worms Blast." In *DirectX 9 Audio Exposed: Interactive Audio Development*, edited by Todd M. Fay, 463–72. Wordware Publishing.
- Marks, Aaron. 2009. *The Complete Guide to Game Audio: For Composers, Musicians, Sound Designers, Game Developers*. 2nd ed. Focal Press.
- McAdams, Stephen, and Bruno L. Giordano. 2009. "The Perception of Musical Timbre." In *The Oxford Handbook of Music Psychology*, edited by Susan Hallam, Ian Cross, and Michael Thaut, 72–80. Oxford University Press.
- Medina-Gray, Elizabeth. 2014. "Meaningful Modular Combinations: Simultaneous Harp and Environmental Music in Two Legend of Zelda Games." In *Music in Video Games: Studying Play*, edited by K. J. Donnelly, William Gibbons, and Neil Lerner, 104–21. Routledge.
- . 2016. "Modularity in Video Game Music." In *Ludomusicology: Approaches to Video Game Music*, edited by Michiel Kamp, Tim Summers, and Mark Sweeney, 53–72. Equinox.
- . 2017. "Musical Dreams and Nightmares: An Analysis of Flower." In *The Routledge Companion to Screen Music and Sound*, edited by Miguel Mera, Ronald Sadoff, and Ben Winters, 562–76. Routledge.
- Miller, Kiri. 2012. *Playing Along: Digital Games, YouTube, and Virtual Performance*. Oxford University Press.
- Moseley, Roger. 2016. *Keys to Play: Music as Ludic Medium from Apollo to Nintendo*. University of California Press. DOI: <https://doi.org/10.1525/luminos.16>.
- Palmer, Caroline, and Carol L. Krumhansl. 1990. "Mental Representations for Musical Meter." *Journal of Experimental Psychology: Human Perception and Performance* 16 (4): 728–41.
- Paul, Leonard J. 2013. "Droppin' Science: Video Game Audio Breakdown." In *Music and Game: Perspectives on a Popular Alliance*, edited by Peter Moormann, 63–80. Springer VS.
- Phillips, Winifred. 2014. *A Composer's Guide to Game Music*. The MIT Press.
- Plomp, R., and W. J. M. Levelt. 1965. "Tonal Consonance and Critical Bandwidth." *The Journal of the Acoustical Society of America* 38 (4): 548–60.
- Reale, Steven B. 2014. "Transcribing Musical Worlds; or, Is L.A. Noire a Music Game?" In *Music in Video Games: Studying Play*, edited by K. J. Donnelly, William Gibbons, and Neil Lerner, 77–103. Routledge.
- . 2016. "Variations on a Theme by a Rogue A.I.: Music, Gameplay, and Storytelling in Portal 2." *SMT-V: Videocast Journal of the Society for Music Theory* 2 (2–3). <https://vimeo.com/173480730> and <https://vimeo.com/191421764>.
- Rogers, Susan E. 2010. "The Influence of Sensory and Cognitive Consonance/Dissonance on Musical Signal Processing." Ph.D. diss., McGill University.

- Sheffield, Brandon. 2008. "Staying In Tune: Richard Jacques On Game Music's Past, Present, And Future." *Gamasutra*. June 16, 2008.
http://www.gamasutra.com/view/feature/3695/staying_in_tune_richard_jacques_.php.
- Snyder, Bob. 2000. *Music and Memory: An Introduction*. The MIT Press.
- . 2009. "Memory for Music." In *The Oxford Handbook of Music Psychology*, edited by Susan Hallam, Ian Cross, and Michael Thaut, 107–17. Oxford University Press.
- Summers, Tim. 2016a. "Analysing Video Game Music: Sources, Methods and a Case Study." In *Ludomusicology: Approaches to Video Game Music*, edited by Michiel Kamp, Tim Summers, and Mark Sweeney, 8–31. Equinox.
- . 2016b. *Understanding Video Game Music*. Cambridge University Press.
- . 2017. "Dimensions of Game Music History." In *The Routledge Companion to Screen Music and Sound*, edited by Miguel Mera, Ronald Sadoff, and Ben Winters, 139–52. Routledge.
- Sweet, Michael. 2015. *Writing Interactive Music for Video Games: A Composer's Guide*. Addison-Wesley.
- Terhardt, Ernst. 1974. "Pitch, Consonance, and Harmony." *The Journal of the Acoustical Society of America* 55 (5): 1061–69.
- Thomas, Chance. 2016. *Composing Music for Games: The Art, Technology and Business of Video Game Scoring*. CRC Press.
- Toiviainen, Petri, and Carol L. Krumhansl. 2003. "Measuring and Modeling Real-Time Responses to Music: The Dynamics of Tonality Induction." *Perception* 32: 741–66.
- Tymoczko, Dmitri. 2011. *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice*. Oxford University Press.
- Vuvan, Dominique T., Jon B. Prince, and Mark A. Schmuckler. 2011. "Probing the Minor Tonal Hierarchy." *Music Perception* 28 (5): 461–72.

Games

- The Legend of Zelda: The Wind Waker*. 2003. Nintendo GameCube. Disc ID: DL-DOL-GZLE-USA. Nintendo.
- Portal 2*. 2011. Steam, Windows PC. Valve Corporation.

Footnotes

1. Sustained academic attention to video game music started gaining momentum most clearly with the publication in 2008 of both a monograph and an edited volume by Karen Collins, and has since then increased continually in scope and output (K. Collins 2008a, 2008b). For a comprehensive list of existing writing on video game music, see the bibliography maintained by the Society for the Study of Sound and Music in Games, at <https://www.ssmg.org/wp/bibliography/>. The relative scarcity of methods for analyzing video game music is especially apparent in comparison with analysis of film music—a relative of video game music—which has had a couple of decades to develop (see Lehman 2017).

[Return to text](#)

2. Interactivity and play are central concerns in a rich and growing body of musicological scholarship that examines music and sound in video games in relation to players' experiences. See, for example, Miller 2012; K. Collins 2013; Cheng 2014; Moseley 2016.

[Return to text](#)

3. Music21 is copyright (c) 2006–2017, Michael Scott Cuthbert and cuthbertLab. Music21 code (excluding content encoded in the corpus) is free and open-source software, licensed under the Lesser GNU Public License (LGPL) or the BSD License. See <http://web.mit.edu/music21/>.

[Return to text](#)

4. For more on video game music in relation to other types of modular and indeterminate music (in particular certain music of the 20th-century avant garde) and aspects of analytical approaches to this other music that may be relevant for analysis of video game music, see [Medina-Gray 2016](#).

[Return to text](#)

5. *Wind Waker* was first released in Japan, and an English-language localized version was released in North America in 2003. In 2013, a remastered version of the game was released for the Wii U as *The Legend of Zelda: The Wind Waker HD*. All examples from *Wind Waker* in this article (including videos, transcriptions, and analyses) are drawn from the 2003 North American version of the game, played on a GameCube console.

[Return to text](#)

6. The names I assign to modules throughout this article are descriptive, and they are not necessarily the same as the names for these modules or cues that appear in any primary sources connected to the game or its creators.

[Return to text](#)

7. A different module plays on “Outset Island” during the first portion of the game, before Link’s sister, Aryll, is kidnapped; this alternate music for “Outset Island” cannot participate in the ocean travel system, and so is not treated here.

[Return to text](#)

8. The speeds of the fades and the length of the silence in this system depend on the speed at which a player is travelling; at normal (quick) speeds, the silence in the middle of this seam is approximately 0.4 seconds long, but the silence is longer if a player crosses this boundary slowly. More accurately, this system appears to rely on a few location-based triggers: one set of triggers determines the volume of “Ocean” from full volume to silence approaching the boundary between ocean and island, another trigger (the arrow shown in Example 1) stops “Ocean” and begins “Outset Island,” and a final set of triggers determines the volume of “Outset Island” from silence to full volume. These triggers are, moreover, tied to the location of the virtual camera, rather than Link’s position in the game’s world. This article’s analyses focus on only the most straightforward situations, in which players sail at relatively quick speeds from the ocean to Outset Island.

[Return to text](#)

9. Nick Collins has adapted the concepts of black box and white box testing from the field of software engineering in order to analyze generative music programs. For an overview of the challenges of using a black box approach, see [N. Collins 2008](#), 240–41.

[Return to text](#)

10. In the instance of the case study from *Portal 2* later in this paper, I was able to access musical files extracted from the game, which aided me in identifying (and transcribing) the modules stored in the game’s memory, but repeated testing through gameplay was still necessary in order to determine how these modules behave in the game.

[Return to text](#)

11. For more on the challenges of accessibility and materials in analyzing video game music, see [Summers 2016a](#).

[Return to text](#)

12. Composers of video game music sometimes refer to these techniques as “horizontal re-sequencing” and “vertical layering” or “vertical remixing” ([Phillips 2014](#), 188–202; [Sweet 2015](#), 143–64).

[Return to text](#)

13. A note on terminology: I use the term “smoothness” because it readily reflects a sense of close and easy combination between modules, especially in the horizontal dimension (e.g., “a smooth connection from one module to the next”), and because other authors already often use terms like “smooth” or “smoothly” when describing these and similar situations in video game music. Words that are normally antonyms of “smoothness” (e.g., roughness, bumpiness), however, do not as readily match the sense that modules do not fit together; “disjunction” seems more applicable here. My use of the term “disjunction” is similar to Hanninen’s use of the same term in relation to boundary-finding and segmentation through sonic criteria ([Hanninen 2012](#)). The use of the terms “smoothness” and “disjunction” in this article is the same as in my earlier work

(Medina-Gray 2014, 2016, 2017).

[Return to text](#)

14. For more on the smoothness aesthetic in video game music (and other popular modular music of the 20th and 21st centuries) see [Medina-Gray 2016](#), 63.

[Return to text](#)

15. For similar suggestions from composers that clear changes in music (that is, disjunction) can productively support changes in the game environment, see: [Hoffert 2007](#), 33–35; [Thomas 2016](#), 69.

[Return to text](#)

16. In a history of game music reception, for example, Summers points out that players’ expectations for repetition and amount of musical material in games has changed over time, and that some types of “abrupt musical changes” that would have been acceptable in the early 1990s were criticized by the mid-2000s ([Summers 2017](#), 148). With regards to genre, William Gibbons has suggested that musical looping and repetition has become part of the conventions and expectations of Japanese Role-Playing Games ([Gibbons 2017](#), 419).

[Return to text](#)

17. For example, Christopher Hasty’s method for segmentation analysis involves examining continuity and discontinuity in various domains in order to identify segments and structure in a piece of music ([Hasty 1981](#)). Dora Hanninen’s overarching theory of music analysis includes attention to sonic criteria in several dimensions as a means of analyzing segmentation among successive as well as simultaneous events ([Hanninen 2012](#)).

[Return to text](#)

18. In Hoffert’s words, his system allows the reader to examine the transition from one musical sequence (that is, a module) to another across a boundary, the effects of which may be “jarring, smooth, or something in between” ([Hoffert 2007](#), 33).

[Return to text](#)

19. Regarding the perceptual integration and segregation of sounds into auditory streams, Albert Bregman points out that “most of the factors that influence the sequential grouping of sounds are those that affect their similarity” ([Bregman 1990](#), 58). Another way to think about smooth sequential seams is that the music after the seam fulfills expectations of similarity and repetition set up by the music before the seam. For more on musical expectation with respect to repetition even in brief periods of exposure, see, for example, [Huron 2006](#), 227–29.

[Return to text](#)

20. Short-term memory is also limited by the number of different elements it can hold, but these elements can group hierarchically to yield varying capacities depending on context ([Snyder 2000](#), 36).

[Return to text](#)

21. For example, Hoffert recommends gauging smoothness by comparing modules’ tempos—whether the music is “slow” or “fast”—and levels of rhythmic activity—whether the music is rhythmically “active” or “passive” ([Hoffert 2007](#), 36–37). In a discussion about writing modules that transition from one larger piece of music to another, Sweet recommends that only one module out of two during an overlap contain rhythmic activity or tempo, to avoid conflict (that is, disjunction) ([Sweet 2015](#), 170). See also [Childs, IV 2007](#), 152.

[Return to text](#)

22. For the probe-tone studies Huron references, see [Palmer and Krumhansl 1990](#). The metric “fit” Huron discusses most directly relates to the situation of sequential seams, but simultaneous seams also involve prevailing metric contexts within which the onsets of both modules either align or not.

[Return to text](#)

23. I hear “Ocean” layers 1 and 2 as supporting the same overall $\frac{4}{4}$ meter regardless of layer 1’s 3+3+2 rhythmic pattern. Alternately, an analyst might highlight the syncopation between the two layers within most measures—3+3+2 against 2+2+2+2—and arrive at a metric analysis of mild smoothness (similar to displacement or grouping dissonance treated below).

[Return to text](#)

24. I actually can tap a tenuous connection across this seam if I focus only on the eighth-note pulse stream: the first couple of onsets in “Outset Island” roughly align with this previous pulse stream, and the new eighth-note pulse stream’s IOIs only differ from the previous pulse stream by 38 ms. However, since the quarter-note pulse streams of both modules—which should also align—are more clearly different, the seam is overall disjunct in terms of meter.

[Return to text](#)

25. Meter (what Krebs calls “metrical consonance”) is here considered to be comprised of nested layers of pulse streams, some streams faster than others, where all pulses in slower streams coincide with pulses in faster streams. Following Krebs, metric dissonance occurs when one or more pulse streams are *not* aligned. Displacement dissonance occurs when two otherwise equivalent pulse streams are displaced from each other by some number of pulses in a shared faster pulse stream (as, for example, in an off-beat syncopation). Grouping dissonance occurs when two pulse streams group the pulses of a shared faster pulse stream by amounts that are not multiples or factors of each other (as, for example, in a 3:2 hemiola).

[Return to text](#)

26. In sequential seams, grouping dissonance may require analytical attention beyond the one-second window after the seam, as necessary.

[Return to text](#)

27. Benadon specifies that, in the repertoire he examines, deviations greater than 50 ms are generally “evident and intentional-sounding” (that is, significant) (Benadon 2009, 6). To my ears, this 50 ms threshold applies to the music in this article as well, but this is an estimate nonetheless.

[Return to text](#)

28. Bregman, for example, references experiments that examine the role of “brightness” in stream segregation, but he points out that this is only one factor within the larger umbrella term “timbre” (Bregman 1990, 93, 96–98). For an introduction to various perceptual aspects of timbre and areas of research in this topic, see McAdams and Giordano 2009.

[Return to text](#)

29. The generalization that unrelated instruments yield strong timbral contrast may not apply to all situations. For example, if the music is recorded or played back at low quality, some of the most significant distinctions between instrumental timbres may fall away.

[Return to text](#)

30. Hoffert notes the keys before and after a sequential seam and measures the distance between those keys around a circle of fifths; the smaller the distance between the two keys, the greater the smoothness (Hoffert 2007, 38–39). Marks recommends writing the various pieces of music in a game in the same or similar keys in order to maintain compatibility between them or tie them together (Marks 2009, 234, 249). Sweet says that composers often write stingers (brief modules) in the same key as other ongoing music to avoid simultaneous conflicts with the underscore (Sweet 2015, 173).

[Return to text](#)

31. Some probe tone experiments involving non-major-key-diatonic contexts also support the idea that repeated tones tend to fit with the preceding context especially well. For example, Vuvan, Prince, and Schmuckler (2011) found that listeners’ ratings of how well a tone “belonged” with a given minor-key context corresponded with which one the three theoretical types of minor scales (natural, harmonic, or melodic) made up that context. Krumhansl, Sandell, and Sergeant (1987) found that a group of listeners with less musical training on average tended to give higher ratings of fit to tones that sounded in a preceding partial-tone-row context (while listeners with more musical training tended to give higher ratings to tones that were absent in the preceding context, presumably because of a more advanced understanding of how tone rows work). Krumhansl and Schmuckler’s (1986) probe tone tests using octatonic scales, however, did *not* result in overall higher ratings for tones in the preceding context, so reoccurrence of pitches is not necessarily a sole or overriding criteria for goodness of fit.

[Return to text](#)

32. Although the D $\flat$ functions as tonic—quite different than the previous C $\sharp$ ’s leading tone function—this new tonal context may not be apparent in the moments immediately after the seam; the pitches make the

connection, and the tonal reorientation comes shortly thereafter.

[Return to text](#)

33. Note that an intuitive focus on key as a measure of sequential smoothness turns out to have some overlap with the method proposed here, since the pitches most likely to occur in a diatonic key are also likely to occur again, and closely related keys share many pitch classes.

[Return to text](#)

34. For an example of a concurrent probe-tone study in which the authors focus on key profiles rather than consonance, see [Toiviainen and Krumhansl 2003](#).

[Return to text](#)

35. This method uses the following steps for determining the intervals created by two simultaneous modules in a given moment: First, reduce each module's pitch content in that moment to a set of pitch classes (with no repetition within the set). Then, remove from each set any pitch classes that appear in both modules (because those pitch classes will not produce any new intervals). For each remaining pitch class in one module, make note of the intervals between that pitch class and each remaining pitch class in the other module.

[Return to text](#)

36. This analytical focus on moment-by-moment musical combination also underlies, for example, Huron's exploration of "tonal fusion" (or lack thereof) through the abundance (or paucity) of various intervals between instrumental parts in J. S. Bach's polyphonic music ([Huron 1991](#)).

[Return to text](#)

37. Plomp and Levelt, for example, have explored consonance and dissonance in relation to critical bandwidth, which depends on the register of tones involved in an interval ([Plomp and Levelt 1965](#)). For one view of the importance of harmonics in perceptions of consonance and dissonance, see [Terhardt 1974](#). Kameoka and Kuriyagawa have provided evidence that intervals larger than an octave may present different levels of consonance and dissonance than do their less-than-octave relatives ([Kameoka and Kuriyagawa 1969](#), 1454). Regarding context, cognitive information might cause a perfect 4th to sound dissonant in certain harmonic contexts, for example. For a brief summary of the distinction between cognitive and psychophysical (sensory) components of consonance and dissonance, see [Rogers 2010](#), 165.

[Return to text](#)

38. This categorization of intervals follows, for example, Ernst Křenek's distinction between "mild" dissonances (major 2nds and minor 7ths) and "sharp" dissonances (minor 2nds and major 7ths) in the context of atonal counterpoint, although Křenek treats the tritone separately as a "neutral" interval ([Křenek 1940](#), 7–8). Some empirical evidence broadly supports this theoretical categorization as well (see: [Rogers 2010](#), 25, 56; [Huron 1994](#), 293–94) although the tritone may more accurately classify as a soft dissonance according to these results, and major 2nds (a soft dissonance) are sometimes reported to be more dissonant than major 7ths (a hard dissonance).

[Return to text](#)

39. It is possible, for example, that macroharmony can play a role in simultaneous situations, as Bjorn Lynne suggests in a self-analysis of his music for the game *Worms Blast*: "I went for the middle ground, writing the music using one scale and a wide variety of chords based on that scale. [. . .] Even if the background music was currently playing an F#sus4 (which doesn't contain a G), if my [simultaneous] motif had a G in it, that was fine, so long as it didn't include a note that was outside of the scale" ([Lynne 2004](#), 466).

[Return to text](#)

40. Note that the right-hand side of the grid for simultaneous seams (in Example 5a) is stretched to incorporate the gradations of pitch-based disjunction, but this is not an indication of relative strength between the two sides. That is, strong disjunction (on the right) is not twice as strong as strong smoothness (on the left).

[Return to text](#)

41. Hanninen argues for a context-dependent approach to weighting of various sonic criteria in segmentation analysis—rather than approaches that use algorithms to determine weighting across many situations—in part because the specific set of weights becomes an analytic and interpretive statement itself ([Hanninen 2012](#), 31–32).

[Return to text](#)

42. When not sailing, only the first layers of “Ocean Intro” and “Ocean” sound, resulting in a different seam.

[Return to text](#)

43. The fact that timbre and pitch both produce mild smoothness rather than strong smoothness may suggest the start of a new subsection in the piece.

[Return to text](#)

44. Stepping away from the modular representation of this music momentarily, the contents of “Ocean Intro” and “Ocean” might also reasonably be represented with a single score, with the contents of “Ocean Intro” first, followed by the contents of “Ocean” within repeat signs. Such a score would be more familiar than the current modular representation, it would still accurately represent the music’s in-game behavior (the progression of this music is automatic in that it requires no input from the player or any other in-game parameters), and the musical continuity from “Ocean Intro” into “Ocean” is obvious enough to support this single-score view. With such a view, using the full five-part method to analyze this single, automatic, pre-determined moment in the music may seem to be uncalled for. However, since the music is indeed modular (some musical content loops while other content does not), the current analysis of the seam between these modules is at least conceptually appropriate. More importantly, since it is precisely the continuity—the smoothness—from one module to the next that allows the music to function as a single larger piece of music, this rather straightforward situation serves as a useful test case for the analytical method’s validity; if the structure is intuitively/obviously conceivable as a single piece of music, then the analysis should reflect this (and it does).

[Return to text](#)

45. Similarly balanced contributions from smoothness and disjunction also govern travel from the ocean to at least two other islands in the game world—Windfall Island and Dragon Roost Island, each with their own associated music—as well as the moves from these islands back to the open ocean. The boundaries between ocean and island become invisible (or rather, inaudible) at night in the game world, when both the ocean and island music are silent.

[Return to text](#)

46. The longest stretch of pitch-based disjunction in this entire simultaneous seam occurs in mm. 53–55, with a sustained F $\sharp$ against G; such disjunction at this point helps to make up for the strong timbral smoothness during this same portion of the seam, so that the two layers can remain at least somewhat distinct.

[Return to text](#)

47. The fact that these two layers intuitively belong to a single piece of music is reflected in my labeling for these modules: “Ocean” layers 1 and 2. In this view of a single piece of music whose layers enter and exit the mix, an analyst might also view these layers with respect to their textural roles (as melody, accompaniment, and so on); for more on issues of texture (and an approach to texture that involves auditory streaming), see, for example, [Duane 2012](#).

[Return to text](#)

48. From a compositional viewpoint, smoothness may be considered a common prerequisite for synchronized layered systems in games. Axel Berndt, for example, has suggested that simultaneous layers in an interactive context “have to harmonize tonally and metrically with each other” so that these parts may fade in and out “without unintended disharmony and rhythmic stumbling” ([Berndt 2009](#), 356).

[Return to text](#)

49. A total of five brief modules—the four in Video Example 2 and one more—consistently accompany conversations throughout *Wind Waker* and so can combine with much more simultaneous music besides just “Ocean Intro” and “Ocean.” The results for pitch-based smoothness vary widely depending on the environment in which the conversation happens.

[Return to text](#)

50. This analysis for meter considers only the moments of the conversation modules’ single main onsets—rather than moments in the modules’ sustained portions—since the onsets contain the only sonic information that might agree or disagree with the simultaneous music’s meter.

[Return to text](#)

51. For examination of yet another music system in *Wind Waker*, see, for example, [Medina-Gray 2014](#).

[Return to text](#)

52. Players also encounter *major* enemies on the ocean, though less frequently; the music that accompanies those situations is different and is not discussed here.

[Return to text](#)

53. Three possible situations are incorporated into these calculations, and each is treated as equally likely: (a) in the boat and sailing: “Ocean Intro”/”Ocean” 1+2 and “Early Combat” 1+2; (b) in the boat but not sailing: “Ocean Intro”/”Ocean” 1 and “Early Combat” 1+2; and (c) out of the boat: “Ocean Intro”/”Ocean” 1 and “Early Combat” 1.

[Return to text](#)

54. Although the ocean’s music is silent at night, the combat music still sounds, so even at nighttime the soundtrack is able to reinforce this critical shift between safety and danger.

[Return to text](#)

55. This analysis takes into account all possible combinations of layered modules and treats these situations as equally likely: (a) Link deals the combat blow and remains in the boat: “Early Combat” 1+2 and “Late Combat” 1+2; (b) the enemy deals the combat blow and knocks Link out of the boat: “Early Combat” 1+2 and “Late Combat” 1; (c) the enemy deals the combat blow while Link is out of the boat: “Early Combat” 1 and “Late Combat” 1.

[Return to text](#)

56. The official soundtrack for *Portal 2* is available for download at <http://www.thinkwithportals.com/music.php>, © 2011 Valve Corporation.

[Return to text](#)

57. Similar musical structures also accompany players’ interactions with gels in other areas of *Portal 2*.

[Return to text](#)

58. I recorded Video Example 4 using the Windows (PC) version of *Portal 2*, accessed through the Steam platform on August 7, 2018.

[Return to text](#)

59. I lean toward the latter, physiological interpretation, since to parallel an air-rushing sound most closely, the real-time volume of the bounce modules would need to correlate with speed rather than height.

[Return to text](#)

Copyright Statement

Copyright © 2019 by the Society for Music Theory. All rights reserved.

[1] Copyrights for individual items published in *Music Theory Online (MTO)* are held by their authors. Items appearing in *MTO* may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may *not* be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of *MTO*.

[2] Any redistributed form of items published in *MTO* must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of *MTO* in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of *MTO*, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may be copied and/or distributed for research purposes only.

Prepared by Sam Reenan, Editorial Assistant