

Manifesto Conchordal

Koichi Takahashi

Music is the drama that unfolds between physical vibration and the human cognitive system that receives it.

Yet conventional music production has confined sound within “grids of symbols”—staff notation and piano rolls—freezing this dynamic relationship into static form. Composition became the act of selecting specific points from an infinitely expansive acoustic space and arranging them along a temporal axis. Sound was reduced to raw material, an object to be dominated by human will.

Conchordal dismantles these existing frameworks. Based on the principle of *Direct Cognitive Coupling*, it reconstructs the mechanisms of human auditory cognition themselves as a landscape within virtual space. Here, sounds are not placed; they are autonomous agents that traverse their environment, interfere with one another, and discover their own harmonies. This represents a paradigm shift from human-controlled “composition” to system-generated “emergence”—a challenge to restore music from static structure to living phenomenon.

The Quest for Harmony: From Pythagoras to the Present

Since antiquity, humanity has sought universal laws governing musical harmony. Pythagoras discovered that intervals could be explained through ratios of string lengths and believed music could be completely described by mathematics. Pythagorean tuning generates all scale degrees through a single operation: multiplying a frequency by three, then folding it back within the octave (2:1) to obtain the perfect fifth—a process using only two prime numbers.

Contrary to Pythagoras’s conviction, however, his tuning system revealed the mathematical imperfection inherent in our world. Twelve iterations of the perfect fifth operation, intended to complete the circle of an octave, yield a pitch slightly higher than the starting note. This inconvenient microtonal discrepancy—the “Pythagorean comma”—prevents pure integer ratios from forming a closed circle, creating an irreconcilable contradiction in tuning. Western music responded to this problem with a compromise: *twelve-tone equal temperament*. By slightly narrowing each of the twelve perfect fifths and distributing the error evenly, the contradiction was resolved, yielding a system

capable of modulation to any key. This was a victory of sorts, yet it simultaneously meant that every interval departed from pure integer ratios—an imperfect solution forcing a choice between the beautiful but unstable just intonation and the stable but slightly impure equal temperament.

In the eighteenth century, Rameau’s *Treatise on Harmony* elevated what had been an accumulation of empirical rules into a theoretical system. Building upon the acoustic discovery of overtones, he attempted to define consonance and dissonance through numerical ratios, establishing the major triad as music’s foundation and articulating the concept of functional harmony. The structure of tension arising from the dissonance of the leading tone’s seventh resolving to the tonic became the foundation of all tonal music from the Baroque through the Romantic era and beyond. Yet Rameau’s theory, while it could explain major chords, left minor chords incompletely accounted for. It never fully answered why humans perceive certain sounds as “pleasing.” Like Pythagoras’s system, Rameau’s theory harbored mathematical incompleteness within it—a factor that, as audiences’ ears grew accustomed over time, led to an ever-expanding catalog of permitted exceptions to harmonic progressions and an endless deferral of resolution to the tonic, shaping the historical trajectory of Western music.

At the dawn of the twentieth century, Arnold Schoenberg’s *Verklärte Nacht* (Transfigured Night) marked one of the ultimate destinations of Romantic tonal music, pushing the post-Wagnerian “delay of resolution to the tonic” to its extreme limit. Subsequently, in the final movement of his String Quartet No. 2, he abandoned resolution to the tonic altogether. In twelve-tone technique, all twelve pitches became equal, and the gravitational field of “return to the tonic” that had long sustained music was lost. This granted composers who followed immense freedom, yet it simultaneously stripped listeners of music’s center of gravity, casting them into a state of weightless drift.

Messiaen, Boulez, and Stockhausen accelerated this trajectory further. In total serialism, not only pitch but duration (rhythm), dynamics, and timbre were governed by rigorous mathematical series. This thoroughgoing rationalism pushed musical language into uncharted territory, yet the works that resulted were so highly abstracted that, paradoxically, listeners often perceived them as “random noise.” Human auditory cognition, with its temporal resolution and pattern recognition capabilities, cannot process minute variations in pitch and mathematical manipulations of rhythm at the same resolution.

As a diametrically opposite approach, Xenakis introduced stochastic processes from physics, and Cage introduced chance operations, both seeking escape from deterministic control. Nevertheless, in both cases, the cognitive reward system—how the human ear craves sound and anticipates resolution—remained relegated behind the structure. Even in the latter half of the twentieth century, audiences continued to seek classical tonal works, and the divide between contemporary

music and its listeners became definitive. With no answer to the fundamental question “What is harmony?” many composers from the late twentieth century onward turned to postmodern eclecticism, partially reviving tonal elements as with Schnittke and Pärt, or quoting and collaging past styles. The horizon of music theory remained shrouded in fog.

Could the key to breaking this impasse lie not in notated symbols but in the physical structure of sound itself? A group of composers who thought so emerged in 1970s France. The spectralists—Grisey, Murail, and others—conceived of sound as physical waveforms, pursuing compositional methods that analyzed and synthesized overtone spectra. Their attempt to construct harmony not from traditional functions or scales but from psychoacoustic facts actually received by the human ear possessed a persuasive new sonority—liberated from conventional tonal grammar yet grounded in physical and physiological reality.

Meanwhile, Toru Takemitsu opened new horizons of sound from an angle distinct from Western modernity. He described his ultimate achievement as a “sea of tonality” and spoke of the concept of *pantonicity*: “a state in which a single tone contains the entire world, and all sounds are organically intertwined”—not mere disorder but chaos harboring cosmic order. This vision, encompassing both tonality and atonality, connects to how harmony in his music is felt not as local functional progression but as a vast “field of resonance.” While rejecting Western functional harmony, he believed in what might be called “the gravitational pull of sound itself.”

Entering the twenty-first century, AI-driven generative models emerged as new agents of creation. Machine learning models trained on massive musical datasets can now automatically generate pieces that sound as if composed by humans. This might appear to be a democratization of creativity, yet a fundamental challenge remains. Current AI primarily extracts and mimics statistically frequent patterns from past data. Consequently, generated music tends to be rehashes or recombinations of previous works, lacking stylistic novelty or the “necessity” of why particular sounds were chosen. Here, clear creative principles such as harmonic theory, counterpoint, or jazz modes are absent—this does not answer the question of “the grounds for sound after the loss of tonality” that the twentieth century exposed.

Direct Cognitive Coupling Through Conchordal

Conchordal is an attempt to dispel the fog that shrouds the origins of music and harmony, and to illuminate a new path.

The essence of Conchordal lies in “generating sound directly upon the human cognitive landscape.” This represents a new paradigm in music production—a concrete implementation with the principle

of *Direct Cognitive Coupling* at its core, directly linking computer-based sound generation processes to the mechanisms of human auditory cognition.

Direct Cognitive Coupling is the principle of grounding sound generation directly in the structure of human auditory cognition without passing through symbolic intermediaries such as notes or scales. Whereas conventional composition undergoes a conversion process from symbol to sound, here the cognitive potential field itself becomes the environment for sound generation.

This principle inherently aims at bidirectional linkage between sound generation and auditory cognition. The current Conchordal implements the first stage: mapping from cognitive model to acoustic environment. In the future, it aims to realize a closed loop—feeding back the listener’s biosignals (brainwaves, heart rate, galvanic skin response, etc.) to the system in real time, so that music transforms according to the listener’s internal state, and that music in turn transforms the listener.

When this closed loop is realized, the distinctions among composer, performer, and audience will dissolve. The one who designs scenarios, the one who physically enters the system, the one who resonates through listening—these will no longer be exclusive roles but gradations of participation. The audience becomes performer; the performer intervenes in composition; the composer dwells within the system. Music is liberated from the fixed, unidirectional flow of “someone creates, someone performs, someone listens,” transforming into a phenomenon of “dwelling together, generating together.”

The Cognitive Landscape of Music

The “landscape” that forms Conchordal’s foundation is a potential field modeled on the characteristics of human auditory cognition. Whereas conventional composition fits sounds into grid-like scales and fixed notes, Conchordal recreates human auditory physiology itself as a virtual environment. The system analyzes the acoustic signal occurring in real time, constructing this internal topography across two dimensions: frequency and time.

The Frequency Axis: Topography of Pitch and Harmony Modeled on the physical structure of the inner ear’s cochlea, a tonotopic map (frequency map) is deployed on a logarithmic scale with base 2. On this axis, the following potentials are calculated in real time based on findings from physiology and cognitive psychology.

Roughness corresponds to amplitude interference within critical bands on the cochlear basilar membrane. Based on the Plomp-Levelt model, it quantifies physiological dissonance caused by beating between closely spaced frequency components. Higher values indicate greater acoustic “coarseness” in that frequency region.

Harmonicity corresponds to phase-locking in neural firing from the cochlear nucleus through the inferior colliculus—the regularity of temporal fine structure. It indicates the degree to which acoustic components are likely to be integrated by the brain as an integer-ratio overtone series, representing the level of perceptual “fusion.”

Consonance is the overall harmonic fitness obtained by integrating Roughness and Harmonicity. It represents the viability of sound at each point on the landscape—the “harmonic” peaks toward which agents should strive.

The Temporal Axis: Topography of Rhythm and Groove Music cognition is deeply linked to neural oscillations in the brain. Conchordal analyzes the temporal evolution of the frequency-axis landscape, extracting in real time a four-layer periodic structure based on brainwave models. Sound agents synchronize with these periods, forming organic rhythms spontaneously rather than through externally imposed meter.

Delta band (~0.5–4 Hz) governs the slow pulse pervading an entire piece—metrical structure and large-scale phrasing (the breath of musical phrases).

Theta band (~4–8 Hz) contributes to articulation and the formation of syllabic-scale groupings—the layer that carves the contours of individual sound events.

Alpha band (~8–13 Hz) shapes pulse and accent patterns within phrases—the intermediate layer forming rhythm’s skeleton.

Beta band (~13–30 Hz) controls groove feel and synchronization precision among multiple sounds (microtiming)—the finest layer where temporal “fluctuation” and “sharpness” coexist.

The landscape thus constructed is based on current findings in psychoacoustics and neuroscience. However, it is not presented as fixed, universal truth. Music perception varies by culture, era, and individual, and the meaning of “consonance” is likewise diverse. Conchordal’s landscape is variable as a parameter, leaving room to incorporate tuning systems from different cultures, individual auditory characteristics, or as-yet-unknown perceptual principles. This is not a system that imposes a specific aesthetic but an open foundation for exploring the direct coupling of cognition and sound.

The Emergence of Acoustic Life

Even if a particular landscape is specified, the music it suggests is not singular. With one parameter setting, Mozartean balance might emerge; with another, punk-rock impulse. The landscape is a space of possibility, and its exploration requires some kind of “inhabitant.” Conchordal adopts artificial life (ALife) agents as these inhabitants.

Individual is the basic unit operating on the landscape. Each Individual is an autonomous acoustic

synthesizer that perceives the potential gradients around it, synchronizes with neural rhythms along the time axis, and survives by expending energy. Individuals seek regions of high consonance as they traverse frequency space; in dissonant regions they rapidly deplete. There is no central conductor. Everything is determined by local perception and reaction.

Population is the acoustic community formed by multiple Individuals. Individuals interfere with one another, differentiating their niches to avoid competition or symbiotically fusing in regions with overtone relationships. Population density dynamically deforms the landscape itself; environment and inhabitants mutually define each other. A Population is not merely the sum of Individuals but possesses emergent properties as a collective.

In this way, musical structures—harmony, melody, rhythm—emerge bottom-up from the local behaviors of Individuals. It is not the composer placing notes but the dynamics of an acoustic ecosystem that generate music.

Scenario: Giving Shape to Time

Music is the act of carving from infinite, uniform physical time a finite duration with beginning and end, and endowing it with structure—tension and release, rise and fall, climax and silence. In Conchordal, harmony and rhythm at the micro level emerge autonomously, but this temporal structure—the arc that makes a work cohere as an experience—is designed by a human creator. The composer acts not as a micromanager placing each note but as a director of the ecosystem.

The tool for this direction is the *Scenario*: a set of instructions that modulates the ecosystem’s environmental conditions along the time axis. The creator specifies not individual sounds but state transitions of the entire system—deforming the landscape’s topography, controlling the generation and extinction of populations, shifting emphasis among rhythm layers. Through these operations, the creator can trace macroscopic trajectories from chaos to order, or from silence to saturation.

The Reach of Direct Cognitive Coupling

The principle of Direct Cognitive Coupling demonstrated by Conchordal is not unique to music. The paradigm of modeling the aesthetic evaluation axes inherent in human sensation and cognition as a landscape, within which creative elements operate autonomously, harbors universality generalizable to all artistic domains.

Visual art: How does the human visual system process the physical phenomena of light wavelength and intensity? Hierarchical structures—color decomposition by retinal cone cells, edge detection in primary visual cortex, color perception in V4, spatial integration in parietal cortex—form the landscape’s foundation. Color balance across hue, brightness, and saturation, and center of gravity

and tension in composition form the potential field; visual-element agents move across the canvas sensing complementary-color contrast (visual roughness) and stability through symmetry.

Literature and narrative: Cognitively modeling harmony in narrative—balance of structure, foreshadowing and payoff, waves of tension and release. Working memory in prefrontal cortex, temporal integration of context via hippocampus, and processing of expectation and deviation through predictive coding form the landscape of narrative cognition. Characters, plot elements, and themes become agents that traverse a landscape of semantic networks and emotional arcs.

Architecture and spatial design: Landscaping comfortable spatial proportions, flow of circulation, and structural stability. Equilibrium through the vestibular system, spatial cognition through place cells in the hippocampus and grid cells in the entorhinal cortex, and the relationship between body and environment through proprioception form the physiological basis of the architectural landscape. Architectural elements become agents, self-organizing according to potentials derived from both physical laws and aesthetic-ergonomic criteria.

In every domain, the core is the same: the evaluative criteria that human senses and cognition have cultivated through long evolution are not superficially imitated as AI training data but incorporated into the system as first principles. Creative outputs emerge not as reproductions of statistically frequent patterns but with cognitive necessity.

Conclusion

This manifesto is a declaration to reclaim the lost gravity of music and to open new horizons of creation.

Since Pythagoras, humanity has sought the principles of harmony in number. Yet neither pure ratios, nor the compromise of equal temperament, nor the abandonment of tonality, nor the mimicry of generative AI could answer the question: “Why must it be this sound?”

We now return to the origin. The principles of harmony do not reside solely in abstract formulas. They dwell in physical phenomena—and in the human ear and brain that receive them. Conchordal is an attempt to embody this fact through engineering. Here, sound is no longer symbol but an entity that behaves autonomously. The acoustic ecosystem and the terrain of human perception are directly bound; true co-creation between machine and human begins. Music returns from fixed recording to living phenomenon.

The name “Conchordal” resonates with multiple meanings: *concord* (harmony), *chordal* (pertaining to harmony), *conch* (evoking the spiral cochlea of the inner ear), and *coral*. Just as a coral reef emerges from the autonomous activity of countless polyps, Conchordal’s music arises from the local

behaviors of individual agents. Its harmony is rooted directly in the spiral, shell-like cochlea—the human organ of hearing.

We present Conchordal to the world as one answer to contemporary music after the dissolution of harmony. It is at once a new art form and one response to fundamental questions: “What is human?” “What is creation?”