Modeling Acoustics in Virtual Environments
Using the Uniform Theory of Diffraction

Nicolas Tsingos[1], Thomas Funkhouser[2], Addy Ngan[2], Ingrid Carlbom[1]

[1] Bell Laboratories*
[2] Princeton University†

Abstract

Realistic modeling of reverberant sound in 3D virtual worlds provides users with important cues for localizing sound sources and understanding spatial properties of the environment. Unfortunately, current geometric acoustic modeling systems do not accurately simulate reverberant sound. Instead, they model only direct transmission and specular reflection, while diffraction is either ignored or modeled through statistical approximation. However, diffraction is important for correct interpretation of acoustic environments, especially when the direct path between sound source and receiver is occluded.

The Uniform Theory of Diffraction (UTD) extends geometrical acoustics with diffraction phenomena: illuminated edges become secondary sources of diffracted rays that in turn may propagate through the environment. In this paper, we propose an efficient way for computing the acoustical effect of diffraction paths using the UTD for deriving secondary diffracted rays and associated diffraction coefficients. Our main contributions are: 1) a beam tracing method for enumerating sequences of diffracting edges efficiently and without aliasing in densely occluded polyhedral environments; 2) a practical approximation to the simulated sound field in which diffraction is considered only in shadow regions; and 3) a real-time auralization system demonstrating that diffraction dramatically improves the quality of spatialized sound in virtual environments.

Keywords: Spatialized Sound, Virtual Environments, Sound Visualization, Uniform Theory of Diffraction, Beam Tracing.

1 Introduction

Realistic simulation of virtual environments has been a major focus of research in interactive computer graphics for decades, dating back to the early flight simulators of the 1960s. Most prior research focused on visualization, while relatively little attention was paid to auralization. However, auditory cues are important in immersive virtual environments, as they combine with visual cues to aid in localization of objects, separation of simultaneous sound signals, and formation of spatial impressions [5] which enhance and reinforce

---

*{tsingos|carlbom}@research.bell-labs.com
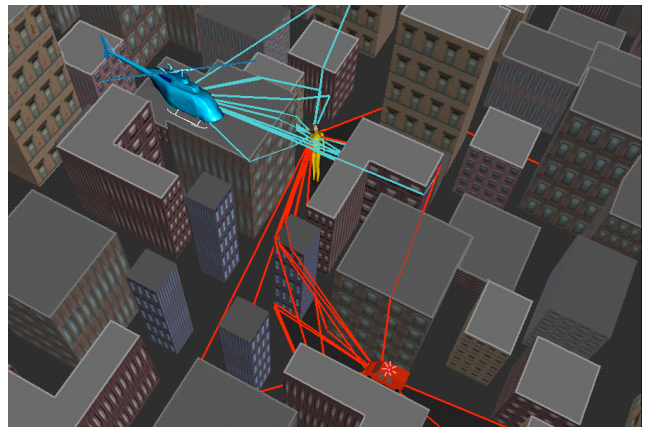†{funk|waingan}@cs.princeton.edu

Figure 1: Early diffracted and reflected sound paths in a city environment where direct sound from sources is occluded. We use the Uniform Theory of Diffraction which considers edges in the environment as sources of new diffracted rays, complementing reflected and transmitted rays.

the visual comprehension of the environment. Experiments have shown that accurate acoustic modeling gives a user a stronger sense of presence in a virtual environment [9], and that high-quality audio enhances perceived visual quality [37].

Current virtual environment systems render audio using geometrical techniques, such as image-sources, ray-tracing or beam-tracing, to compute early propagation paths between sound sources and listener while late reverberation is usually modeled using statistical techniques [12, 13, 34]. Unfortunately, they fail to spatialize sound realistically because they do not accurately account for diffraction. Our goal is to compute early geometrical propagation paths, including diffraction effects, that can be used for real-time auralization in such a system.

Diffraction is a form of scattering by obstacles whose size is of the same order of magnitude as the wavelength. It is a fundamental mode of sound propagation, particularly in building interiors and cities where the direct path between a sound source and a receiver is often occluded. For example, consider the training simulation scenario shown in Figure 1, in which a pedestrian must respond to a helicopter and a fire engine. Failure to simulate sound diffraction could be "disastrous," since the person cannot see either sound source, and the direct path from both sound sources is blocked by tall buildings. In such situations, it is important that the virtual environment system simulates diffraction correctly because people localize sounds by interpreting echoes according to their delays relative to the first arriving wavefront [5, 30]. Moreover, if diffraction is omitted from a virtual environment simulation, the user may experience abrupt changes in spatialized sounds as he/she turns a corner and the sound source disappears from the line of sight. For instance, consider walking down the hallway of your office building and having the sound disappear after you pass each open door.

Such abrupt changes in a simulation would introduce a mismatch with our real world experiences, which would result in "negative training" or at least confuse users.

In this paper, we propose an efficient way for computing the acoustical effect of early reflection and diffraction paths according to the Uniform Theory of Diffraction [21, 24, 26]. Specifically, we make three contributions. First, we describe a beam tracing method, for enumerating sequences of diffracting edges efficiently and without aliasing in densely occluded polyhedral environments. Second, we propose an approximation to simulated sound fields suitable for immersive virtual environments in which diffraction is computed only in shadow regions. Finally, we describe a real-time auralization system that produces spatialized sound with early diffraction, transmission, and specular reflection during interactive walkthroughs of complex environments.

Our experimental results demonstrate that (1) beam tracing is an efficient and aliasing-free way to find diffraction sequences in densely occluded environments, (2) it is possible to construct early diffracting propagation paths and spatialize sounds in real-time, and (3) diffraction greatly improves the quality of spatialized sounds in immersive virtual environments.

2   Background and Related Work

There are currently three major approximation theories for diffraction problems in polyhedral environments: (1) the Huygens-Fresnel diffraction theory, (2) boundary integral representations using the Helmoltz-Kirchoff integral theorem, and (3) the Uniform Theory of Diffraction.

Huygens' principle [17] predicts that every point on a wavefront can be regarded as the source of a secondary spherical wavelet. The wavefield is defined at each point by the superposition of these wavelets, which extend in all directions, including shadow regions. Fresnel supplemented Huygens' theory by adding interference between the wavelets to treat diffraction [6]. He also subdivided space between the source and the receiver into concentric ellipsoids with frequency-dependent radii: the Fresnel ellipsoids. By modeling diffraction effects as a loss in signal intensity, Bertoni [4] and Tsingos and Gascuel [43] use Fresnel ellipsoids to determine relevant obstacles at any given frequency. By replacing the binary geometrical visibility by an extended visibility term between 0 and 1, they achieve frequency-dependent sound "muffling." This technique removes abrupt cuts in the simulated sound, producing a more pleasing experience. However, it fails to capture the temporal aspect of diffraction since new propagation paths are not introduced. While this approximation is not usually a concern for electromagnetic wave propagation, it is an important issue for acoustics.

Analytic expressions give time-domain diffraction filters for sequences of finite wedges based on a discrete Huygens interpretation [27, 38]. In this case, the edges are discretized into secondary point sources whose contributions must be summed to obtain the diffracted field. Such models prove very accurate for low order diffraction and have recently been used to assess audibility of diffraction in the case of a simple stage house [41]. However, it is unclear if the method can be applied in real-time since the edges must be discretized into a large number of point sources to compute the diffraction filters.

The Helmoltz-Kirchoff integral theorem provides a formalization of the Huygens-Fresnel principle [6, 10]. It expresses the scattered field at any point in space as a function of the field on the surface of the diffracting objects. Mathematically, it can be expressed as a surface integral and solved by numerical methods such as Boundary Element Methods (BEM) [16, 18] that discretize surfaces into patches. BEM allow for very accurate treatment of sound diffraction. But, they are far too compute intensive for interactive sound rendering over the whole audio spectrum and are mainly used

for low frequencies (below 150Hz). In some cases the integral can be solved analytically [35], such as for height fields or periodic surfaces. However, neither of these cases usually applies to architectural models.

The Uniform Theory of Diffraction (UTD) [21, 24, 26] incorporates diffraction into the ray theory of light. The UTD treats an infinite wedge as a secondary source of diffracted waves that in turn can be reflected and diffracted before reaching the receiver. For a given point source and point receiver location, the diffraction of a wave over an infinite wedge is represented by a *single* ray whose contribution to the wave field is attenuated by a complex valued diffraction coefficient [24] (see Appendix A). For any sequence of diffracting edges, the ray follows the path satisfying Fermat's principle: if the propagation medium is homogeneous, the ray follows the shortest path from the source to the receiver, stabbing the diffracting edges. The UTD is a high frequency approximation and applies to infinite wedges, when the source and listener remain far from diffracting surfaces (compared to the wavelength). To date, the UTD has been applied successfully in several types of off-line simulations, including acoustical diffraction over solitary wedges [20], lighting effects in room-sized scenes [2], and radio frequency propagation in buildings and cities [33, 22]. For acoustic waves, the method has been validated down to 150Hz for a small combination of diffracting wedges [20, 38]. Validation of the approach for more complex situations has not yet been achieved.
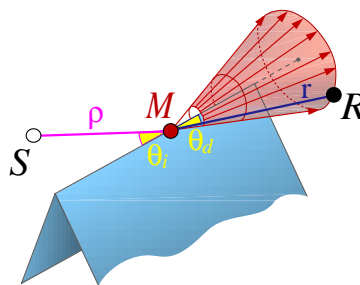

Figure 2: According to the UTD, an incoming ray ρ gives rise to a cone of diffracted rays, where the aperture angle θ_d of the cone is equal to the angle θ_i between the incident ray and the edge (the axis of the cone is the edge). For a given receiver location, a single ray describes the diffracted field.

Of these three approaches, the UTD is the most promising for spatializing sound in interactive virtual environments, as it integrates well into a geometrical framework, is physically-based, and provides satisfying results for most of the audio spectrum (for early diffraction orders).

The main computational challenge in using the UTD into real-time auralization systems is the efficient enumeration of significant early diffraction paths. Although many algorithms exist to find approximate solutions [29], they are either too inefficient or prone to aliasing. For instance, Aveneau [2] enumerated all permutations of polyhedral edges within the first few Fresnel ellipsoids, which is not practical for sound simulations in large environments. Rajkumar et al. [33] extended a ray tracing algorithm to broadcast rays in all directions for each edge "intersection." Similarly, Fortune et al. [11, 22] and Stephenson [36] described a beam tracing approach in which propagation due to each edge diffraction was approximated by a set of beams emanating from point sources at sampled positions along the diffracting edge. These latter two approaches approximate the set of potential diffraction paths by discrete sampling. Thus they are prone to aliasing, which would cause noticeable artifacts in an interactive sound system. Prior methods provide neither interactive response times nor guarantee finding all significant propagation paths.
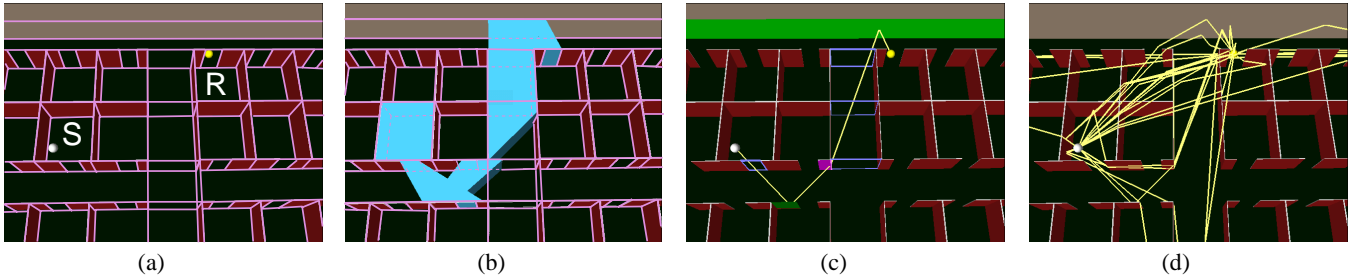
<div align="center">(a)   (b)   (c)   (d)</div>

Figure 3: Overview of our process: (a) Virtual environment (office cubicles) with source S, receiver R, and spatial subdivision marked in pink. (b) Sample reflected and diffracted beam (cyan) containing the receiver. (c) Path generated for the corresponding sequence of faces (green), portals (purple), wedges (magenta). (d) The procedure repeated for all beams containing R.

In this paper, we describe real-time methods for simulating early sound reverberation with the Uniform Theory of Diffraction in large virtual environments. We address two main issues of such a system: (1) enumerating significant propagation paths efficiently, and (2) computing an approximation to the diffracted field that can be updated at interactive rates. Details are provided in the following sections.

3 Enumerating Propagation Paths

According to the UTD, an acoustic wave incident upon an edge between two non-coplanar surfaces forms a diffracted wave that propagates in a cone shaped pattern of rays from the intersected part of the edge, as shown in Figure 2. Our challenge is to represent the propagation pattern of these rays as they traverse free space, transmit through obstacles, and reflect off surfaces.

Our approach is based on object-precision beam tracing [15]. The motivation for this approach is to exploit the spatial coherence in propagation paths while avoiding the aliasing artifacts of sampling diffraction edges. In contrast to ray tracing, beam tracing works with object-precision polyhedral volumes that support well-defined intersections with diffracting edges. Aliasing resulting from the intersection of infinitely thin rays with infinitely thin edges is thus eliminated [33]. In contrast to brute-force enumeration of all edge permutations [2], beam tracing provides an effective method for pruning the search based on the feasibility of stabbing lines [39]. As a result, beam tracing finds every propagation path up to a specified termination criteria without undersampling errors. Moreover, beam tracing algorithms are practical for specular reflection in densely occluded virtual environments [12], and can be readily incorporated into interactive virtual environments systems [13].

In the following two subsections, we focus on the challenges of tracing beams and constructing propagation paths with diffraction. Unfortunately, as beams emanating from a source and diffracting over an edge are traced along subsequent sequences of reflections and transmissions, they can become quite complex, bounded by quadric surfaces due to triple-edge (EEE) events. Rather than representing these scattering patterns exactly [8, 40], we conservatively over-estimate the space of rays diffracting over an edge with a polyhedral approximation. We compensate for this approximation later by checking each propagation path to determine if it lies in the over-estimating part of the beam, in which case it is discarded. Since diffraction patterns are approximated conservatively and tightly with simple polyhedra, and checking propagation paths is quick, the whole process is much faster than computing the exact propagation pattern directly.

3.1 Beam Construction and Tracing

The goal is to enumerate the significant permutations of diffractions, specular reflections, and transmissions along which a sound wave can travel from a given source location. The algorithm must be conservative, so that no significant propagation paths are missed. But, it should not be too over-estimating, so that the second stage of our process becomes over-burdened with construction of infeasible propagation paths. Finally, to enable efficient checking of propagation paths, our algorithm must not only construct a beam containing the region of space reachable by each propagation sequence, but it must also encode potential blockers (or equivalently "portals").

We incrementally compute beams starting from a source by traversing the cell-face and face-edge adjacency graph of a polyhedral cell complex, as in [19, 1, 12, 13, 39] (see Figure 3). Starting in the cell containing the source with a beam representing the entire cell, we iteratively visit adjacent cells in priority order, considering different permutations of transmission, specular reflection, *and diffraction* resulting from the faces and edges on the boundary of the "current" cell. As each new cell C is visited, the current beam B is updated such that it contains all potential propagation paths along the current traversal sequence. We identify diffracting edges ε on the boundary of C as the ones: (1) intersected by B and (2) shared by two faces F_1 and F_2 on the boundary of C that are either non-coplanar or have different acoustic properties (e.g., F_1 is transparent and F_2 is opaque). For each such edge, we construct a new beam containing potential diffraction paths and begin tracing it through all adjacent cells. We also construct and trace beams for transparent and reflecting surfaces, as in [12]. All sequences and their corresponding beams are logged in a *beam tree* data structure [12, 15], which can be queried later to determine the set of propagation paths reaching a specific receiver location.

Each beam emanating from a diffracting edge and passing through or reflecting off a sequence of cell boundaries is represented conservatively by the intersection of two cones and a polytope (see Figure 4). The two cones are constructed with axes along the diffracting edge, with apexes at the two endpoints of the beam-edge intersection. Their interior angles are derived from the equal angle constraints at these endpoints, as dictated by the Uniform Theory of Diffraction. The polytope bounds the set of lines emanating from the diffracting edge and stabbing the traversed sequence of convex cell boundaries with a constant number of opposing planes, as described in [40]. This representation allows every beam traced through a sequence of arbitrarily oriented faces to keep a bounded complexity, which is important for both computational efficiency and memory utilization. Accordingly, using the adjacency information in the cell complex, each beam is updated incrementally in expected-case constant time.

Although there are generally exponentially many distinct sequences of diffraction through a 3D polyhedral scene, we expect a large number of them to be psychoacoustically insignificant (the amplitude of diffracted contributions quickly drops with the diffraction order), and thus beams are traced in priority order [28], either during an off-line precomputation (as in [12]) or in real-time using multiple asynchronous processes (as in [13]).
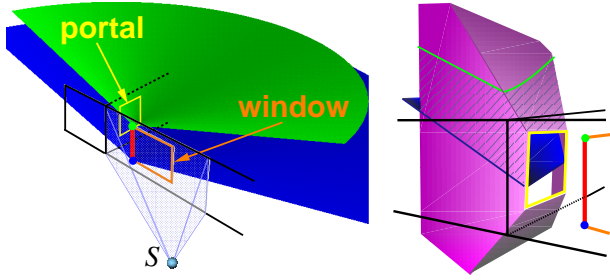
Figure 4: Possible diffraction paths are conservatively bounded by the intersection of two cones and a polytope. Left: a beam incident on the edge (red) of a window and the two diffraction cones (shown in blue and green only in the halfspace behind the window) computed at the endpoints of the beam-edge intersection. Right: a close-up view of the intersection (hatched) between the two cones and the polytope resulting from stepping through the next portal (yellow).

3.2 Path Construction and Validation

Once beams are traced from a sound source, we construct a unique propagation path for each beam containing a receiver location. The geometry of the path determines the delay, amplitude, and directivity of the sound wave traveling along the path from the source to the receiver.

According to the UTD, the wave field resulting from a diffraction can be approximated by a piecewise-linear propagation path – i.e., the shortest among all possible paths from the source to the receiver stabbing the faces and edges in the sequence. In order to construct this path for a given sequence of beams, we find the points of intersection for every reflecting face and diffracting edge. The intersections with specularly reflecting faces are uniquely determined by the locations of the source, receiver and diffraction points. Thus, the problem is reduced to finding the locations of the diffraction points, P_i ($i = 1...n$) (see Figure 5). At each of these points, the path satisfies a simple "unfolding property" (see Figure 2): the angle (θ_i) at which the path enters the edge must be the same as the angle (θ_d) at which it leaves [14]. Thus, for each potential path, we solve a non-linear system of n equations expressing equal angle constraints at the diffracting edges:

$$
\begin{cases}
\overrightarrow{P_1 S} \cdot \overrightarrow{E_1} & = & \overrightarrow{P_1 P_2} \cdot (-\overrightarrow{E_1}) \\
\overrightarrow{P_2 P_1} \cdot \overrightarrow{E_2} & = & \overrightarrow{P_2 P_3} \cdot (-\overrightarrow{E_2}) \\
& \vdots & \\
\overrightarrow{P_n P_{n-1}} \cdot \overrightarrow{E_n} & = & \overrightarrow{P_n R} \cdot (-\overrightarrow{E_n})
\end{cases}
\tag{1}
$$

where S is the source point, R is the receiver point, $\overrightarrow{E_i}$ is the normalized direction vector of the ith diffracting edge, and $\overrightarrow{P_{i+1} P_i}$ is a normalized direction vector between two adjacent points in the shortest path. To incorporate specular reflections in this equation, $\overrightarrow{E_i}$ and $\overrightarrow{P_{i+1} P_i}$ are transformed by successive mirroring operators accounting for the sequence of specularly reflecting faces up to the ith diffraction.

Parameterizing the edges, $P_i = O_i + t_i \overrightarrow{E_i}$ (where O_i is a reference point on edge i), the system of equations (1) is rewritten in terms of n unknowns t_i and solved within a specified tolerance using a non-linear system solving scheme. We use a locally convergent Newton scheme [32], with the middle of the edges as the initial estimate for the diffraction points. Since the equation satisfied by any diffraction point only depends on the previous and next
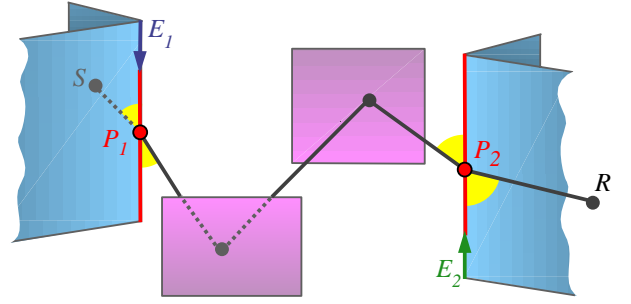

Figure 5: A single propagation path comprising a diffraction, two specular reflections, and another diffraction. The two diffraction points (P_i) are determined by equal angle constraints at the corresponding edges (E_i).

diffraction points in the sequence, the Jacobian matrix is tridiagonal and is easily evaluated analytically. Thus, every Newton iteration is performed in time $O(n)$ where n is the number of unknowns (i.e., edges). We found this method faster than the recursive construction proposed by Aveneau [3].

Once the diffraction points are found, we validate whether the resulting path intersects every cell boundary in the sequence (to compensate for the fact that the beams are conservatively approximate). If not, the path in the over-estimating part of the beam is discarded. Otherwise, it contributes to an *impulse response* used for spatializing sounds [23, 25] (see Appendix A).

The proposed conservative beam tracing and path construction enumerate all sequences of diffracting edges (without aliasing) up to a specified termination criterion, while most acoustically infeasible sequences of edges and faces are culled already during beam tracing.

4 Shadow Region Approximation

Our beam tracing technique provides a method for finding diffraction paths efficiently and without aliasing. However, contrary to specular reflections or transmissions, diffraction introduces a scattering in all directions around the wedge, which results in a combinatorial explosion of the number of beams to consider, even for moderately complex scenes. In this section we introduce an approximation to reduce the spatial extent of diffraction beams, while preserving a good modeling of the diffracted field. This approximation enables us to achieve interactive auralization in large environments.

Recent psychoacoustic tests in the context of a simple stage house model [41] show that diffractions can be perceived in illuminated regions where direct and reflected contributions from a source also reach the listener. However, in this case, we conjecture that diffraction does not modify the main acoustic cues already carried by the direct and reflected contributions since the amplitude of the diffracted field is usually much weaker than direct or even reflected contributions [31] (p.500). On the other hand, we note that diffraction into shadow regions is crucial for typical virtual worlds as it provides the primary mode of propagation to most of the environment. Thus, we introduce an approximation in which *the contribution of diffraction is considered only in shadow regions*.

Accordingly, our current on-line implementation allows for adding an extra halfspace to the polytope representing each diffraction beam so that it tightly, yet conservatively, bounds the shadow region of each diffraction.

However, discarding the diffracted field in the illuminated region of a wedge introduces a discontinuity at the shadow boundary, as the direct field is abruptly replaced by the diffracted field. This is due to the fact that the UTD diffracted field is defined to ensure that the *sum* of the direct and diffracted fields is continuous for any
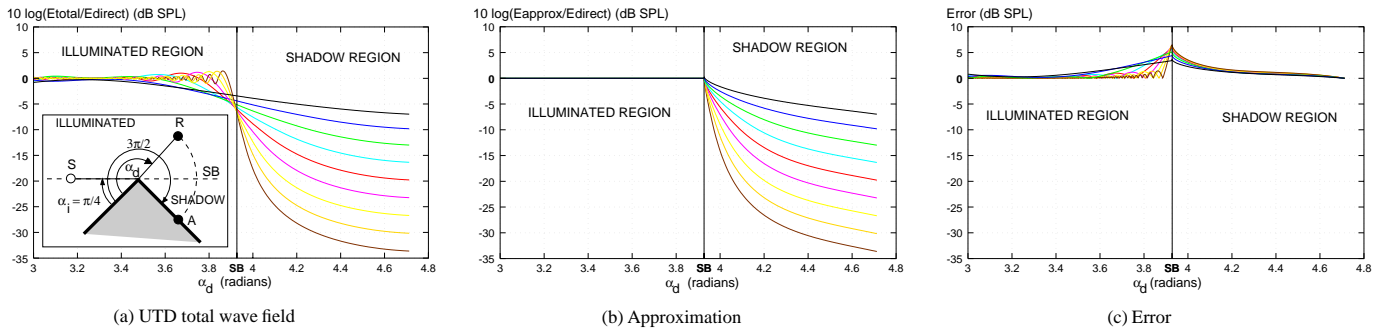
Figure 6: Plots of (a) the UTD total wave field, (b) our approximation, and (c) the error as a function of diffraction angle (α_d), as the receiver rotates around the edge, for a single diffracting wedge (inset). Each plot shows several curves corresponding to the sound pressure level (SPL) for the center frequencies of octave bands ranging from 100 Hz (top) to 24kHz (bottom). Our approximation culls the diffracted field contribution in the illuminated region of the wedge but still closely matches the original UTD field.

receiver location (see Figure 6(a)), while both fields, independently, are discontinuous at shadow boundaries.

To preserve continuity at shadow boundaries, we normalize the diffracted field as predicted by the UTD, so that it is C^0 continuous with the direct field at the shadow boundary, *SB*. We define the normalized diffracted field at the receiver location R as:

$$\mathcal{E}'_{\mathrm{diffracted}}(R) = \mathcal{E}^{\mathrm{SB}}_{\mathrm{incident}}(R)\mathcal{E}_{\mathrm{diffracted}}(R)/\mathcal{E}^{\mathrm{SB}}_{\mathrm{diffracted}}(R), \quad (2)$$

where $\mathcal{E}^{\mathrm{SB}}_{\mathrm{incident}}(R)$ and $\mathcal{E}^{\mathrm{SB}}_{\mathrm{diffracted}}(R)$ are the incident and diffracted fields when the receiver R is rotated to lie on the shadow boundary SB (at the same distance from the edge).

This modified expression scales the diffracted field equally for all directions around the edge, unnecessarily modifying the original UTD field away from the shadow boundary. Hence, our new approximated diffracted field is derived by interpolating between the expression given by equation (2) and the original UTD expression of the diffracted field [24] (see Appendix A) as the receiver further moves inside the shadow region (between SB and A in Figure 6(a)). Since the expressions are complex-valued, care must be taken in the interpolation: *argument* (i.e., phase) and *modulus* must be independently interpolated to give a new complex value.

Figure 6 shows a comparison of the total wave field as predicted by the UTD and our approximated wave field for the situation shown in Figure 6(a). Although there are differences in the vicinity of the shadow boundary, most properties of the original UTD diffracted field are captured by the approximation: (1) the edge is still the source of the diffracted contribution and path delays are not modified by our approximation, (2) the field is continuous, so no audible artifact is heard when crossing the shadow boundary, (3) the field amplitude is independent of frequency at the shadow boundary, and (4) it decays faster as frequency increases and tends toward the actual value of the UTD diffracted field as the receiver moves away from the shadow boundary. As a result, we conjecture that the spatialized sound produced by our on-line system provides many of the significant cues useful for localization of objects, separation of signals, and comprehension of space in an immersive virtual environment.

5 Simulation Results

The 3D data structures and algorithms described in the preceding sections are implemented in C++ and run both on SGI/Irix and PC/Windows computers. We integrated them into a prototype system that allows a user to move through a virtual environment interactively, while images and spatialized audio are rendered in real-time according to the user's simulated position.

To test if our beam tracing approach is practical for modeling diffraction in typical virtual environments, and to evaluate the benefits of incorporating diffraction into real-time auralization, we ran a series of tests computing propagation paths both with and without diffraction. During each test, we used a 3D model with 1,762 polygons representing one floor of a building (see Figure 8). For simplicity, we assumed that every polygon in the 3D model was 80% reflective and acoustically opaque (no transmission). Before each test, we traced 50,000 beams in breadth-first order from a stationary sound source (located at the white dot in Figure 8) and stored them in a beam tree data structure. Then, as a receiver moved at three inch increments along a hallway (the long vertical one on the right side of each image in Figure 8), we computed propagation paths from source to receiver, updated an impulse response, and auralized spatialized sound in real-time. All the tests were run on a Silicon Graphics Onyx2 workstation using two 195MHz R10000 processors, one of which was dedicated to software convolution of audio signals.

The test sequence was executed three times, once for each of the following beam tracing constraints:

1. **Specular reflection only:** We traced 50,000 beams along paths of specular reflection, with no diffraction. The results represent the state-of-the-art prior to this paper [12, 13].
2. **Diffraction only:** We traced 50,000 beams along paths of diffraction (around silhouette edges into shadow regions), with no specular reflections.
3. **Both specular reflection and diffraction:** We traced 50,000 beams along paths representing arbitrary permutations of specular reflection and diffraction (into shadow regions).

Figure 7 shows plots with the number of propagation paths (the top plot) and the power of the impulse responses (the bottom plot) for each receiver location during the three tests.[1]

From these plots, we confirm that specular reflection alone is not adequate to produce realistic spatialized sound in typical virtual environments. The red curves in Figure 7 show that the number of propagation paths and the power in the corresponding impulse responses varied dramatically with small changes in receiver location. This effect is easily understood by examining images of the beams and power distributions shown in Figure 8(a-d). Note the pattern of thin beams zig-zagging across the hallways in the top-left image. As the receiver walks along the test trajectory, s/he moves in and out of these distinct specular reflection beams, leading to sharp discontinuities in the computed early reverberations. Even worse, there are several locations where no specular reflection paths reached the

---

[1] Power is computed as $10 \cdot \log \sum_{i=1}^{n} a_i^2$ where n is the number of propagation paths and a_i the amplitude along the ith path.

(a) Number of propagation paths.

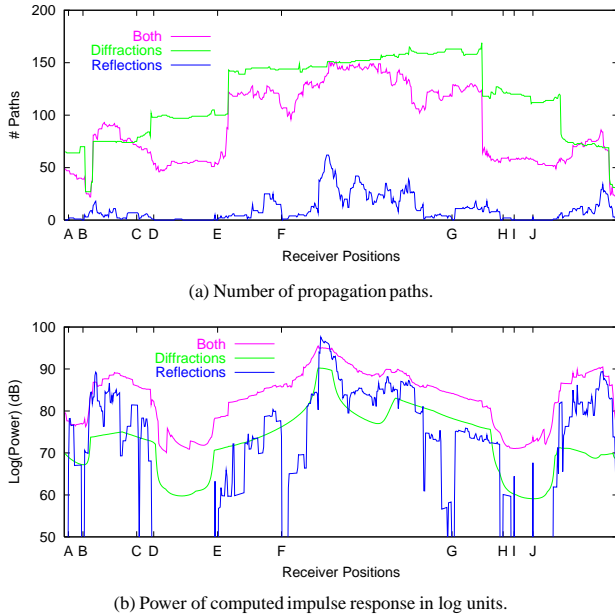(b) Power of computed impulse response in log units.

Figure 7: Plots showing the number of propagation paths (top) and power of the corresponding impulse responses (bottom) computed for every receiver position during tests with specular reflection only (red), diffraction only (blue), and both specular reflection and diffraction (purple). Note how adding diffraction gives a smoothly varying sound level as the receiver moves.

receiver, and thus the power of the auralized sound drops to zero for short periods. These locations, which are marked with capital letters on the horizontal axis of the plots in Figure 7, correspond to the visible "holes" in the beam coverage in Figure 8(a). They are particularly troublesome to users during a walkthrough, as suddenly appearing "dead-zones" clearly fail to match our real-world acoustical experiences.

In contrast, we note that tracing beams only along paths of diffraction leads to smoothly varying reverberations (the blue curves in Figure 7). The reason can be seen in Figure 8(e-h). Diffraction beams tend to cover larger volumes of space than specular beams; and, in our test, they collectively cover all reachable parts of the 3D environment (the light gray regions in the middle of the model correspond to elevator and wiring shafts unreachable by sound). Even though we traced diffraction beams only into shadow regions, our approximation produced a smoothly varying impulse response (see the power map in Figure 8(f)) because direct paths were replaced by diffracting ones with equal amplitude at shadow boundaries as the receiver moved past open doors.

Finally, the test with both specular reflections and diffractions (the purple curves in Figure 7) shows that the power varies quite smoothly, while early reflections and diffractions due to the environment are clearly evident. The improvement in reverberation can be seen clearly by examining the echograms (temporal plot of scattered energy reaching the receiver) shown in the rightmost images of Figure 8. Each pulse in these plots corresponds to a propagation path (shown in yellow in the third column of Figure 8). Note that the echogram measured with both specular reflections and diffractions contains not only the shortest (diffracted) path from the source (the left-most spike), but it also has many high-power early contributions not found in the other tests because they are reflections of previously diffracted waves, or diffractions of previously reflected waves. These contributions combine with the earliest arriving sound wave to provide the dominant acoustical cues.

We also gathered computational statistics during the three tests (Table 1). Column 2 shows the rate (in beams/second) at which beams are traced from the stationary source location in each test. Column 3 shows the rate (in paths/second) at which propagation sequences are computed and processed to form propagation paths to the moving receiver location, including calculation of reflection and diffraction coefficients. The next three columns show the average number of transmissions through transparent cell boundaries (Trans), specular reflections (Refl), and diffractions (Diff) along the computed paths. Finally, the right-most column (Update Time(s)) shows the time (in seconds) required to update the impulse response for each new receiver location in the three tests. Based on these results, we conclude that tracing propagation paths with both diffractions and specular reflections is quite practical for interactive virtual environment applications. Although diffraction increases the time required to trace beams and construct propagation paths (by almost $2\times$), the system still updates impulse responses at interactive time steps (every 49ms) with our method.

Test	Compute Rates		Path Statistics			Update
Name	Beams/s	Paths/s	Trans	Refl	Diff	Time (s)
Specular	6,305	4,289	4.8	3.9	0.0	0.002
Diffract	3,173	1,190	6.8	0.0	5.2	0.163
Both	3,778	2,943	4.2	1.8	1.7	0.049

Table 1: Beam tracing and path generation statistics.

Figure 9 shows visualizations of our results for different application domains. The left-most pair of images shows the power of sound reaching different parts of a city from a siren located on top of a building. In this case, diffraction due to edges of large buildings is the dominant acoustical effect. The second set of images explores the acoustical variations of different seats in the Opéra de la Bastille theater in Paris. There, diffraction over the lip of the orchestra pit provides the primary means for sound to reach the audience, and the slanted balconies are responsible for significant occlusion and diffraction effects. Finally, the rightmost pair of images shows how spatialized sound with diffraction can be used to enhance an interactive video game, as sound diffracting through non-axial obstacles and over walls helps players find each other.

6 Discussion

Our current beam tracing implementation is practical only for densely-occluded and coarsely detailed 3D models, since beams get fragmented in scenes with many free-space cell boundaries [12, 39]. However, since this class of models contains many types of interesting acoustical environments (e.g., buildings and cities), our system is useful for the proposed application domain. To work well for sound simulations in detailed 3D models, our beam tracing algorithm would have to be enhanced, possibly by extending Fortune's topological beam tracing method [11] to work for edge sources.

The UTD is an approximate model of sound propagation, valid mostly for high frequencies and infinite wedges. However, our technique for enumerating propagation sequences can be used in combination with other theories, such as [38], in the context of off-line simulations. The accuracy would be improved, especially at low frequencies and near-field from the wedges. Also, as mentioned in [41], the shortest paths constructed by our technique would still be useful for efficient auralization even if the diffraction coefficients (or filters) are derived from another theory.
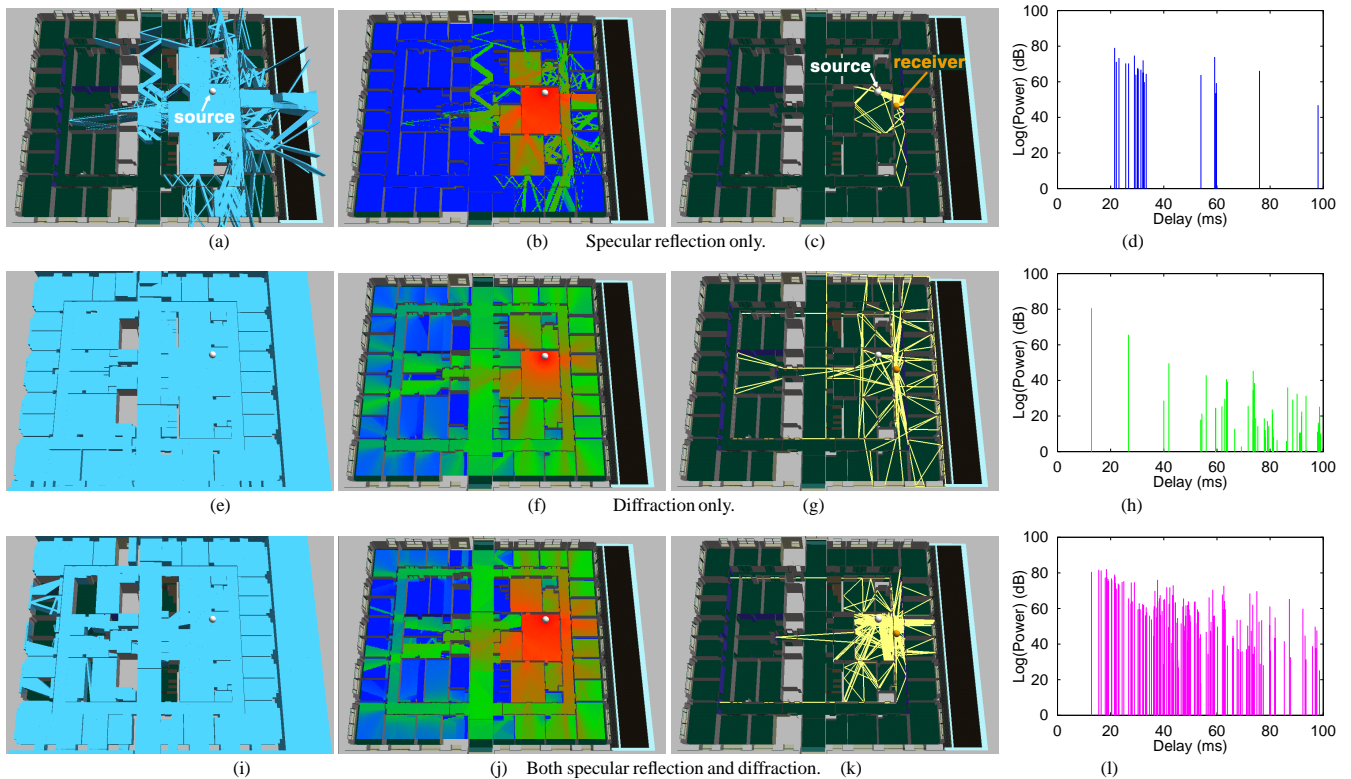
Figure 8: Images depicting results of our experiment with 50,000 beams for specular reflection only (top row), diffraction only (middle row), and both specular reflection and diffraction (bottom row). The left-most column shows beams (cyan) traced from a source point. The next column shows the corresponding power plots (red represents higher power and green lower power). The third column shows computed propagation paths to receiver point. Finally, the right-most column shows the corresponding echograms. Note how diffraction beams fill the entire space and how diffraction combines with reflection to produce a more complete acoustical impression.

7 Conclusion and Future Work

In this paper we introduce an efficient technique for incorporating diffraction effects in interactive audio simulations for virtual environments. Relying upon the Uniform Theory of Diffraction, we describe a beam-tracing approach to construct propagation paths with diffraction, and we introduce a practical approximation to the diffracted field in shadow regions. This is the first instance where a realistic, physically-based, diffraction model is used to produce sound at interactive rates in complex virtual environments. By simulating diffraction, we remove the disturbing "cuts" in the audio that occur when a sound source is occluded by an acoustically opaque surface, and make it possible to localize occluded sound sources. Based on our initial experiences, we conclude that it is possible to compute diffraction paths in real-time and that diffraction dramatically improves the realism and quality of the audio experience.

This research suggests several directions for further study. First, evaluation of simulation results by comparison to measured data is essential. Towards this end, we have recently built a "room" whose walls have known acoustical bidirectional reflectance distribution functions and that allows addition of panels to create interesting geometries, including diffracting panels. We are using this room to verify our simulations and evaluate different approximations. Perceptual assessments, which are probably the most important for virtual environments, will also be conducted.

Second, application of the proposed methods to problems beyond acoustics is a promising topic for future work. We are currently investigating hybrid beam tracing and path tracing approaches to global illumination in which coarsely detailed beams are used to guide the sampling and intersection of paths in a Monte Carlo lighting simulation. This would also be useful to efficiently simulate diffuse surface reflection for acoustic simulations [7, 42]. Other potential applications include motion planning, transmitter power prediction, fire simulation, and traffic analysis.

Finally, perhaps the most interesting topic for future work is the study of the inter-play between visual and auditory stimuli in human perception of 3D environments. Accurate simulations of both sound and light in an interactive system may provide a useful tool for perceptual psychologists to investigate this important question.

Acknowledgments

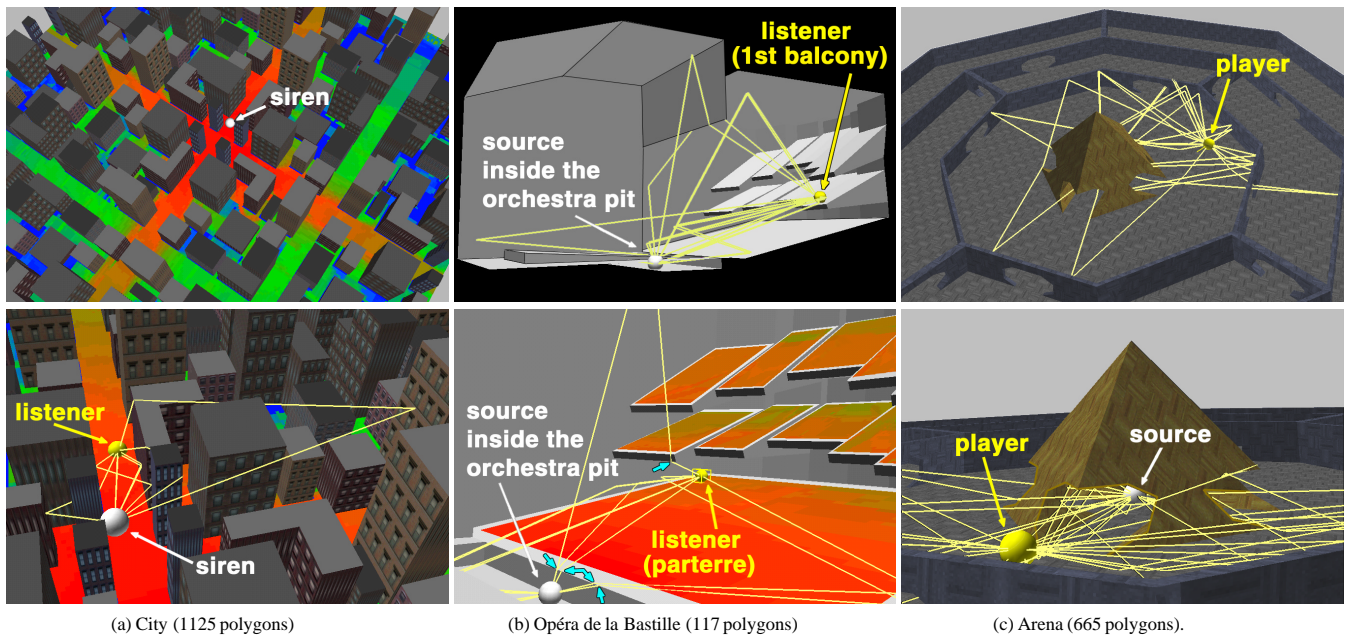(a) City (1125 polygons)   (b) Opéra de la Bastille (117 polygons)   (c) Arena (665 polygons).

Figure 9: Visualizations of sound simulations for different applications. (a) Acoustic power coverage map for a siren in a city environment. (b) Study of early propagation paths for a source located in the orchestra pit of an opera house. Note the diffracted paths over the lip of the pit and balconies (cyan arrows). (c) Diffracted sound paths allow players of a video game to localize the source hidden under the pyramid.

References

[1] J.M. Airey, J.H. Rohlf, and F.P. Brooks, Jr. Towards image realism with interactive update rates in complex virtual building environments. In Rich Riesenfeld and Carlo Séquin, editors, *Computer Graphics (1990 Symposium on Interactive 3D Graphics)*, pages 41–50, March 1990.

[2] L. Aveneau and M. Meriaux. Rendering polygonal scenes with diffraction account. *Seventh International Conference in Central Europe on Computer Graphics and Visualization (Winter School on Computer Graphics)*, February 1999.

[3] L. Aveneau, Y. Pousset, R. Vauzelle, and M. Mériaux. Development and evaluations of physical and computer optimizations for the 3d utd model. *AP2000 Millennium Conference on Antennas & Propagation (poster)*, April 2000.

[4] H.L. Bertoni. Coverage prediction for mobile radio systems operating in the 800/900 MHz frequency range. *IEEE Transactions on Vehicular Technology (Special Issue on Mobile Radio Propagation)*, 37(1), February 1988.

[5] J. Blauert. *Spatial Hearing : The Psychophysics of Human Sound Localization*. M.I.T. Press, Cambridge, MA, 1983.

[6] M. Born and E. Wolf. *Principles of Optics*. 7th ed., Pergamon Press, 1999.

[7] B.-I. L. Dalenbäck. Room acoustic prediction based on a unified treatment of diffuse and specular reflection. *J. of the Acoustical Soc. of America*, 100:899–909, 1996.

[8] G. Drettakis. *Structured Sampling and Reconstruction of Illumination for Image Synthesis*. PhD thesis, University of Toronto, January 1994.

[9] N.I. Durlach and A.S. Mavor. Virtual reality scientific and technological challenges. National Research Council Report, National Academy Press, 1995.

[10] P. Filippi, D. Habault, J.P. Lefevre, and A. Bergassoli. *Acoustics, basic physics, theory and methods*. Academic Press, 1999.

[11] S.J. Fortune. Topological beam tracing. In *Proc. 15th ACM Symposium on Computational Geometry*, pages 59–68, 1999.

[12] T. Funkhouser, I. Carlbom, G. Elko, G. Pingali, M. Sondhi, and J. West. A beam tracing approach to acoustic modeling for interactive virtual environments. *ACM Computer Graphics, Proc. SIGGRAPH98*, pages 21–32, July 1998.

[13] T. Funkhouser, P. Min, and I. Carlbom. Real-time acoustic modeling for distributed virtual environments. *ACM Computer Graphics, Proc. SIGGRAPH99*, pages 365–374, August 1999.

[14] J. Goodman and J. O'Rourke, editors. *Handbook of Discrete and Computational Geometry*. CRC Press, 1997.

[15] P. Heckbert and P. Hanrahan. Beam tracing polygonal objects. *Computer Graphics (SIGGRAPH84)*, 18(3):119–127, July 1984.

[16] D.C. Hothersall, S.N. Chandler-Wilde, and M.N. Hajmirzae. Efficiency of single noise barriers. *J. of Sound and Vibration*, 146(2):303–322, 1991.

[17] C. Huygens. *Traité de la Lumiere*. London, Macmillan & Co., 1912.

[18] P. Jean. A variational approach for the study of outdoor sound propagation and application to railway noise. *J. of Sound and Vibration*, 212(2):275–294, 1998.

[19] C. B. Jones. A new approach to the `hidden line' problem. *Computer Journal*, 14(3):232–237, August 1971.

[20] T. Kawai. Sound diffraction by a many sided barrier or pillar. *J. of Sound and Vibration*, 79(2):229–242, 1981.

[21] J.B. Keller. Geometrical theory of diffraction. *J. of the Optical Society of America*, 52(2):116–130, 1962.

[22] S.C. Kim, B. Guarino, T. Willis, V. Erceg, S. Fortune, R. Valenzuela, L. Thomas, J. Ling, and J. Moore. Radio propagation measurements and prediction using three-dimensional ray tracing in urban environments at 908 MHz and 1.9 GHz. *IEEE Trans. on Vehicular Technology*, 48:931–946, 1999.

[23] M. Kleiner, B.I. Dalenbäk, and P. Svensson. Auralization - an overview. *J. of the Audio Engineering Society*, 41(11):861–875, November 1993.

[24] R.G. Kouyoumjian and P.H. Pathak. A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface. *Proc. of IEEE*, 62:1448–1461, November 1974.

[25] H. Lehnert and J. Blauert. Principles of binaural room simulation. *Applied Acoustics*, 36:259–291, 1992.

[26] D.A. McNamara, C.W.I. Pistorius, and J.A.G. Malherbe. *Introduction to the Uniform Geometrical Theory of Diffraction*. Artech House, 1990.

[27] H. Medwin, E. Childs, and G. Jebsen. Impulse studies of double diffraction: A discrete huygens interpretation. *J. Acoust. Soc. Am.*, 72:1005–1013, 1982.

[28] P. Min and T. Funkhouser. Priority-driven acoustic modeling for virtual environments. *Proc. Eurographics'2000*, 2000.

[29] J. S. B. Mitchell. Geometric shortest paths and network optimization. In Jörg-Rüdiger Sack and Jorge Urrutia, editors, *Handbook of Computational Geometry*. Elsevier Science Publishers B.V. North-Holland, Amsterdam, 1998.

[30] B. C.J. Moore. *An introduction to the psychology of hearing*. Academic Press, 4th ed., 1997.

[31] A.D. Pierce. *Acoustics. An introduction to its physical principles and applications*. 3rd ed., American Institute of Physics, 1984.

[32] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C, 2nd ed*. Cambridge University Press, New York, 1992.

[33] A. Rajkumar, B.F. Naylor, F. Feisullin, and L. Rogers. Predicting RF coverage in large environments using ray-beam tracing and partitioning tree represented geometry. *Wireless Networks*, 2(2):143–154, 1996.

[34] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen. Creating interactive virtual acoustic environments. *J. of the Audio Engineering Society*, 47(9):675–705, September 1999.

[35] J. Stam. Diffraction shaders. *ACM Computer Graphics, Proc. SIGGRAPH99*, pages 101–110, August 1999.

[36] U. Stephenson and U. Kristiansen. Pyramidal beam tracing and time dependent radiosity. *15th International Congress on Acoustics*, pages 657–660, June 1995.

[37] R. L. Storms. *Auditory-Visual Cross-Modal Perception Phenomena*. PhD thesis, Naval Postgraduate School, Monterey, California, September 1998.

[38] U. P. Svensson, R. I. Fred and J. Vanderkooy Analytic secondary source model of edge diffraction impulse responses. *J. of the Acoustical Society of America*, 106:2331–2344, 1999.

[39] S. Teller. Computing the antiumbra cast by an area light source. *Computer Graphics (SIGGRAPH92)*, 26(2):139–148, 1992.

[40] S. Teller. *Visibility Computations in Densely Occuded Polyhedral Environments*. PhD thesis, Computer Science Div., Univ. of California, Berkeley, 1992.

[41] R. Torres, P. Svensson and M. Kleiner. Computation of edge diffraction for more accurate room acoustics auralization. *J. of the Acoustical Society of America*, 109:600–610, 2001.

[42] R. Torres. Studies of Edge Diffraction and Scattering: Applications to Room acoustics and Auralization. PhD thesis, Dept. of Applied Acoustics, Chalmers University of Technology, Sweden, 2000.

[43] N. Tsingos and J.-D. Gascuel. Soundtracks for computer animation: sound rendering in dynamic environments with occlusions. *Proceedings of Graphics Interface'97*, pages 9–16, May 1997.

552

A Auralizing the wedge diffracted field

According to the UTD, the acoustic pressure field diffracted by a wedge can be expressed in terms of the incident field on the edge, $\mathcal{E}_{\mathrm{incident}}(M)$, as:

$$\mathcal{E}_{\mathrm{diffracted}}(R) \quad = \mathcal{E}_{\mathrm{incident}}(M) \, D \, A(r, \rho) \, e^{-ikr}, \qquad (3)$$

where R is the receiver location, M is the diffraction point (see Figure 2), $A(r, \rho) = \sqrt{\rho r/(\rho + r)}$ is a scalar distance attenuation term along the propagation path, the complex exponential e^{-ikr} represents phase variation along the diffracted path, $k = 2\pi/\lambda$ is the wave number (λ is the wavelength). Equation (3) is applied successively for every diffracting wedge and multiplied by attenuations due to reflections, transmissions, and path length to form a contribution to the impulse response for every propagation path.
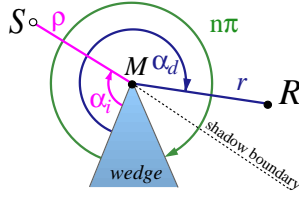

Figure 10: Notations for the UTD diffraction coefficient.

D is the complex-valued UTD diffraction coefficient [24, 26] accounting for amplitude and phase changes due to diffraction:

$$
\begin{aligned}
D(n, k, \rho, r, \theta_i, \alpha_i, \alpha_d) = & -\frac{e^{-i\frac{\pi}{4}}}{2n\sqrt{2k\pi}\sin\theta_i} \\
& \Big[\tan^{-1}\left(\tfrac{\pi+(\alpha_d-\alpha_i)}{2n}\right) F\left(kLa^+(\alpha_d-\alpha_i)\right) \\
& + \tan^{-1}\left(\tfrac{\pi-(\alpha_d-\alpha_i)}{2n}\right) F\left(kLa^-(\alpha_d-\alpha_i)\right) \\
& + \Big\{ \tan^{-1}\left(\tfrac{\pi+(\alpha_d+\alpha_i)}{2n}\right) F\left(kLa^+(\alpha_d+\alpha_i)\right) \\
& + \tan^{-1}\left(\tfrac{\pi-(\alpha_d+\alpha_i)}{2n}\right) F\left(kLa^-(\alpha_d+\alpha_i)\right) \Big\} \Big],
\end{aligned}
\qquad (4)
$$

where (see also Figure 10 and Figure 2):

$$F(X) = 2i\sqrt{X}e^{iX} \int_{\sqrt{X}}^{+\infty} e^{-i\tau^2} d\tau, \qquad (5)$$

$$L = \frac{\rho r}{\rho + r}\sin^2\theta_i, \qquad (6)$$

$$a^{\pm}(\beta) = 2\cos^2\left(\frac{2\pi n N^{\pm} - \beta}{2}\right), \qquad (7)$$

$N^{\pm}$ is the integer that satisfies more closely the relations:

$$2\pi n N^+ - \beta = \pi \quad \text{and} \quad 2\pi n N^- - \beta = -\pi \qquad (8)$$

Several approximations exist in the related literature, useful for implementation of Eq. 4. In particular, relations (8) reduce to:

$$
\begin{aligned}
N^+ &= \begin{cases} 0 & \text{for } \beta \leq \pi(n-1) \\ 1 & \text{for } \beta > \pi(n-1) \end{cases}, \\[1em]
N^- &= \begin{cases} -1 & \text{for } \beta < \pi(1-n) \\ 0 & \text{for } \pi(1-n) \leq \beta \leq \pi(1+n) \\ 1 & \text{for } \beta > \pi(1+n) \end{cases},
\end{aligned}
\qquad (9)
$$

and Kawai [20] gives an approximate rational expression for the integral in Eq. (5):

$$\text{for } X < 0.8 : F(X) = \sqrt{\pi X}\left(1 - \frac{\sqrt{X}}{0.7\sqrt{X}+1.2}\right) e^{i\frac{\pi}{4}\sqrt{\frac{X}{X+1.4}}}$$

$$\text{for } X \geq 0.8 : F(X) = \left(1 - \frac{0.8}{(X+1.25)^2}\right) e^{i\frac{\pi}{4}\sqrt{\frac{X}{X+1.4}}}$$

$$(10)$$

Cotangent terms in Equation (4) are still singular at a reflection or shadow boundary and cannot be evaluated numerically at these boundaries. However, in the vicinity of such a boundary we can express the terms $\alpha_i \pm \alpha_d$ as $\beta = 2\pi n N^{\pm} \mp (\pi - \varepsilon)$. The coefficient is continuous and its value can be computed using [24]:

$$\tan^{-1}\left(\tfrac{\pi \pm \beta}{2n}\right) F\left(kLa^{\pm}(\beta)\right) \simeq$$

$$ne^{-i\pi/4}\left(\sqrt{2\pi kL}\,\mathrm{sgn}(\varepsilon) - 2kL\varepsilon e^{-i\pi/4}\right),$$

where $\mathrm{sgn}(\varepsilon) = 1$ if $\varepsilon > 0$ and -1 otherwise.

In order to render the virtual sound field, we compute a digital filter [23, 25], with which audio signals emanating from the source can be convolved to produce a spatialized audio signal with reverberation. For high quality auralization, this filter is computed using complex values in Fourier frequency space at the desired sampling rate resolution. For interactive applications, fewer frequency bands can be considered, depending on how much processing power is available. The modulus of the complex field for the center frequency of each frequency band can be used to *re-equalize* the source signal. For more information on the signal processing involved in auralization, please refer to [23, 25, 34]