OPEN ACCESS Freely available online
D. Ackermann, F. Brinkmann, and S. Weinzierl,
"A Database with Directivities of Musical Instruments,"
J. Audio Eng. Soc., vol. 72, no. 3, pp. 170–179 (2024 Mar.).
https://doi.org/10.17743/jaes.2022.0128.

A Database with Directivities of Musical Instruments

DAVID ACKERMANN,* FABIAN BRINKMANN, *AES Associate Member* **AND STEFAN WEINZIERL**
(david.ackermann@tu-berlin.de) (fabian.brinkmann@tu-berlin.de) (stefan.weinzierl@tu-berlin.de)

Audio Communication Group, Technische Universität Berlin, Germany

This article presents a database of recordings and radiation patterns of individual notes for 41 modern and historical musical instruments, measured with a 32-channel spherical microphone array in anechoic conditions. In addition, directivities averaged in one-third-octave bands have been calculated for each instrument, which are suitable for use in acoustic simulation and auralization. The data are provided in Spatially Oriented Format for Acoustics. Spatial upsampling of the directivities was performed based on spherical spline interpolation and converted to OpenDAFF and Generic Loudspeaker Library formats for use in room acoustic and electro-acoustic simulation software. For this purpose, a method is presented for how these directivities can be referenced to a specific microphone position in order to achieve a physically correct auralization without coloration. The data is available under the CC BY-NC 4.0 license.

0 INTRODUCTION

Studies of the sound radiation characteristics of the human voice date back to the late 1930s [1], and studies of the directivity of musical instruments began 30 years later (summarized in [2, Chap. 4]). While these early measurements were often made with a single microphone moved around the source, the radiation patterns of acoustic sound sources such as loudspeakers, singers, or musical instruments are now usually measured with the source at the center of an enclosing microphone array in anechoic conditions. In the following, the authors give an overview of existing databases on the directivity of musical instruments. An overview of the directivity of the human voice is beyond the scope of this article, and interested readers are referred to the reviews by Abe [3] and Pörschmann [4].

The directivity of a selection of eight orchestral instruments was measured acoustically using a 64-channel spherical microphone array [5]. For each instrument, the directivity of 12 fundamental tones and their associated six to 16 harmonics were quantified. Using correlation analysis, the radiation patterns of the derived partials were analyzed for each instrument to determine their interdependence. In a similar way, 14 instruments of a symphony orchestra were measured using a 22-channel microphone array [6], and the directivity of each instrument was characterized using either octave or one-third-octave band averaging. The directivity data itself, however, was not made accessible with the two published studies.

In a more recent approach, the radiation patterns of 14 musical instruments was measured with a combination of single-capture recordings with an arc-shaped microphone array and repeated measurements with musicians rotated stepwise under this array, providing a total of 2,522 sampling points on a sphere. The inherent problem of the repeatability of the notes played was partially corrected by using the spectral distribution of the notes at a reference microphone, which allowed the correction of deviations in timbre but not those caused by the movements of the musicians. These directivities are freely available in one-third-octave resolution [7], while the recorded individual notes are not published as such.

One of the most comprehensive public databases to date includes 41 contemporary and historical instruments, including a soprano. This resource used 32 microphones for its measurements and features recordings of individual notes within the playable range of each instrument and directivities derived from the stationary portions of these notes [8, 9].

Based on these measurements,[1] the aim of this work was to provide the directivities of musical instruments in an

*To whom correspondence should be addressed: e-mail: david.ackermann@tu-berlin.de

[1]The data are taken from version 2 of the repository: https://doi.org/10.14279/depositonce-5861.2.

open and standardized format, including both the acoustic measurements, the results of the subsequent processing, and important metadata, such as the exact position of the microphone capsules, tuning frequency of the instrument, and pitch for which the directivity is valid. This should facilitate the exchange and use of such data by the scientific and general acoustics community. To facilitate this, the recently standardized Spatially Oriented Format for Acoustics (SOFA) convention `FreeFieldDirectivityTF` was used [10, 11], but the data are also provided in OpenDAFF [12] and Generic Loudspeaker Library (GLL) [13] formats for use in room acoustic simulation software. For each instrument, the database contains the single-note recordings [calibrated to an absolute sound pressure and equalized for the microphone array transfer function (TF)], extracted single-tone directivities for each partial tone, and one-third-octave band averaged directivities and corresponding finite impulse responses (FIRs) [9].

1 METHODS

The directivity of 40 modern and historical string, woodwind, brass, and percussion instruments and one soprano singer was measured using a 32-channel full-spherical microphone array in an anechoic chamber. Table 1 gives an overview of the instruments contained in the data base including the range of musical notes for which recordings are available. The measurements were presented in an earlier publication [8]. The following sections detail the recording of the original database and highlight the extended and improved processing of the data in this work.

1.1 Data Representation and Format

The directivity of electro-acoustic sound sources, such as loudspeakers and microphones, can be described relatively easily by using TFs in the frequency domain or by FIR filters in the time domain for each direction of radiation. In contrast the directivity of natural sound sources, such as human speakers, singers and musical instruments, is more complex to describe, because the directivity of a musical instrument depends not only on the frequency but also on the note being played, and sometimes on the fingering, so that the same note can have multiple radiation patterns. The different openings for the sound radiation of a bassoon, which can be open or closed depending on the note played and fingering used, are shown in Fig. 1 as an example. The note-dependent directivity has also been shown to produce audible differences in sound [14].

To ensure maximum flexibility, the `Freefield DirectivityTF` convention represents directivity data as complex TFs at arbitrary frequencies. This allows FIR filters of artificial sound sources but also multi-channel acoustic recordings of a musical instrument (see SEC. 1.3.1) to be stored as complex spectra with linear frequency resolution. In addition the directivity patterns of natural sound sources can be stored for arbitrary, not necessarily equidistant, frequencies such as the fundamental frequency and corresponding overtones (see SEC. 1.3.2), or as one-third-octave band averaged data (see SEC. 1.3.3). This representation is mainly used in geometric acoustic simulation.

Table 1. List of modern (m) and historical (h) instruments whose individual directivities are available in the database in the indicated pitch range for the playing dynamics pianissimo (*pp*) and fortissimo (*ff*).

Group	Instrument	Era	Scale *pp*	Scale *ff*
Strings	Violin	m	G3–G6	G3–G6
	Violin	h	G3–C7	G3–C7
	Viola	m	C3–F7	C3–F7
	Viola	h	C3–C6	C3–C6
	Cello	m	C2–G6	C2–G6
	Cello	h	C2–D5	C2–D5
	Double bass	m	E1–E5	E1–D#5
	Double bass	h	E1–E3	E1–E3
	Acoustic guitar	m	C2–B5	E2–B5
	Double action harp	m	C1–F4	C1–F4
Woodwinds	Oboe	m	A#3–A6	A#3–F#6
	Classic oboe	h	C4–D#6	C4–D#6
	Romantic oboe	h	B3–C6	D4–G5
	English horn	m	E3–G#5	E3–G#5
	Clarinet	m	D3–A#6	D3–A#6
	Clarinet	h	D3–F6	D3–F6
	Bass clarinet	m	A#1–D#5	A#1–D#5
	Bassoon	m	A#1–E5	A#1–E5
	Classic bassoon	h	A#1–C5	A#1–C5
	Baroque bassoon	h	A1–F#4	A1–F#4
	Contrabassoon	m	A#0–D#3	A#0–F3
	Dulcian	h	C2–G4	C2–G4
	Alto saxophone	m	C#2–C#5	C#2–C#5
	Tenor saxophone	m	B1–A4	B1–A4
	Flute	m	B3–A#6	B3–D7
	Transverse flute	h	D4–F#6	D4–A#6
	Keyed flute	h	C#4–A6	C#4–A6
Brass	Trumpet	m	F#3–F6	F#3–F6
	Natural trumpet	h	D3–D6	D3–D6
	French horn	m	D2–F#5	D2–G5
	Natural horn	h	A1–B4	A1–B4
	Basset horn	h	F2–A#5	F2–A#5
	Tenor trombone	m	G1–F5	G1–F5
	Bass trombone	m	E1–F4	E1–F4
	Alto trombone	h	D2–D#5	C#2–D#5
	Bass trombone	m	B1–B4	C1–B4
	Tenor trombone	h	E1–D5	E1–D5
	Tuba	m	G#0–C5	F0–E5
Percussion	Timpani	m	D2	D2
	Pedal timpani	m	D2	D2
	Soprano	···	G3–G#5	G3–G#5

For a complete description of natural sound source directivities, information about the instrument/singer and way of playing is needed in addition to the measurement setup. The SOFA standard allows these data to be stored as metadata and uniquely assigned for easy handling and data exchange. An overview of the available metadata can be found in SEC. 2.

1.2 Measurement Setup

The instruments were recorded with a fully spherical lightweight microphone array in the anechoic chamber of the Technische Universität Berlin with a room volume of approximately 1,070 m^3 and lower cut-off frequency of $f_c =$

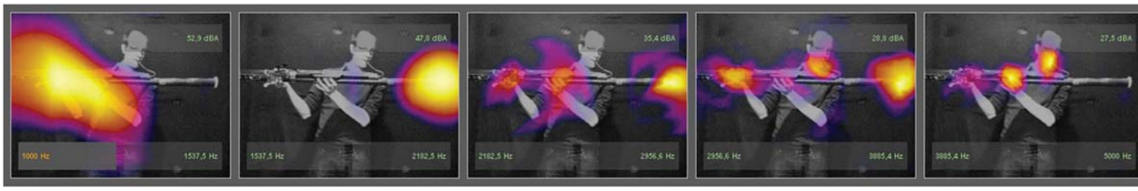

Fig. 1. The sound radiation of the bassoon for the note E#3 in five different frequency bands, from 1 to 5 kHz, recorded simultaneously with an acoustic camera. Bright spots indicate high sound radiation from the instrument. Distributed, frequency-dependent sound radiation becomes visible from the open tone holes on the body, bell, and mouthpiece. Depending on the note played and fingering used, these tone holes can be open or closed, leading to different sound radiation, even for the same frequency [15].

63 Hz. Thirty-two Sennheiser KE4-211-2 electret capsules of the microphone array were located at the vertices of a pentakis dodecahedron with a radius of $r = 2.1$ m, a central angle of $\Theta = 37.4^\circ$, and an arc length of $s = 1.37$ m between the individual microphones.

A height-adjustable chair was used to position the musicians with their instruments so that the estimated acoustic center of each instrument was as close as possible to the center of the microphone array, and the musicians faced the positive x-axis. For extended instruments with secondary sound sources, such as the bassoon with its bell and multiple spaced tone holes, the geometric center of the sound-emitting surface of the instrument has been defined as the acoustic center. The musicians were asked to play all playable notes of their instruments in both extremes of the dynamic strength, i.e., in pianissimo (*pp*) and fortissimo (*ff*) (for precise instructions see [16]), without moving the instrument; the instruments were not fixed by any technical means during the recording.

The exact capsule positions of the array are given in spherical coordinates, i.e., in azimuth ($\phi = 0^\circ$ pointing in positive x-direction, $\phi = 90^\circ$ pointing in positive y-direction), colatitude ($\theta = 0^\circ$ pointing in positive x-direction, $\theta = 90^\circ$ pointing in positive z-direction), and distance (in meters) are included in the metadata of the SOFA files (cf. SEC. 2). The recordings of the single notes of each instrument were made with 24-bit resolution and a sampling frequency of $f_s = 44.1$ kHz. The recordings were calibrated—i.e., a digital amplitude of 1 corresponds to a pressure of 1 Pascal or a sound pressure level of $L_p = 94$ dB – and compensated for the frequency response of the microphone array and capsules. The exact measurement setup and a detailed description of the calibration procedure can be found in [8].

1.3 Processing

The calibrated and equalized single-note recordings in *ff* and *pp* provided as 32-channel WAV files, as published in [9], form the basis for the processing described below.

1.3.1 Recordings

The real-valued single-note recordings $x_q[n]$ of even length N are available at discrete times $n \in \{0, 1, ..., N - 1\}$ and for $Q = 32$ channels of the spherical microphone array with $q \in \{1, 2, ..., Q\}$. Because the `FreeFieldDirectivityTF` convention requires frequency data, the recordings were Fourier transformed

$$X_q(k) = \sum_{n=0}^{N-1} x_q[n] \mathrm{e}^{-\mathrm{i}2\pi \frac{k}{N} n}, \tag{1}$$

with $\mathrm{i}^2 = -1$ being the imaginary unit, and saved in the SOFA files as complex-valued single-sided spectra $X_{\mathrm{S},\,q}(k)$ of length $N/2 + 1$

$$X_{\mathrm{S},q}(k) = \begin{cases} X_q(k), & \text{if } k = 0 \\ 2 \cdot X_q(k), & \text{if } 0 < k < \frac{N}{2} \\ X_q(k), & \text{if } k = \frac{N}{2}. \end{cases} \tag{2}$$

Note that reconstructing the time domain recordings from the published data thus requires the reconstruction of the both-sided spectrum of length N

$$X_q(k) = \begin{cases} X_{\mathrm{S},q}(k), & \text{if } k = 0 \\ \frac{1}{2} \cdot X_{\mathrm{S},q}(k), & \text{if } 0 < k < \frac{N}{2} \\ X_{\mathrm{S},q}(k), & \text{if } k = \frac{N}{2} \\ \frac{1}{2} \cdot X^*_{\mathrm{S},q}(N-k), & \text{if } k > \frac{N}{2} \end{cases}, \tag{3}$$

with $(\cdot)^*$ denoting the complex conjugate, before applying the inverse Fourier transform (see [17] for details).

1.3.2 Single Tone Directivity Data

To determine the directivity of the musical instruments, the authors used the stationary part of the single note recordings. For all instruments producing stationary parts, this was manually windowed by visual inspection, resulting in durations between 200 and 2,104 ms, with a median duration of 630 ms. For the acoustic guitar and harp, a quasi-stationary part was defined as the part between the decay time and release time as estimated with the Timbre Toolbox [18]. For the transient timpani signals, the entire recording was used, from the onset to the transition into the noise floor.

The directivities were estimated in two steps. First the fundamental frequency f_0 and frequency of the overtones f_i in hertz were identified with $i \in \{1, 2, ..., I\}$ and I being the highest identifiable overtone. Secondly the energy at these frequencies was estimated. Both steps were based on the magnitude response $|X_q(k)|$ and ignored the phase information. Determining the phase of a directivity TF would require a reference signal. While for loudspeaker directivities the input signal can be used as a reference, there is no useful reference for a musical instrument, because it is an extended sound source with a note-dependent acoustical

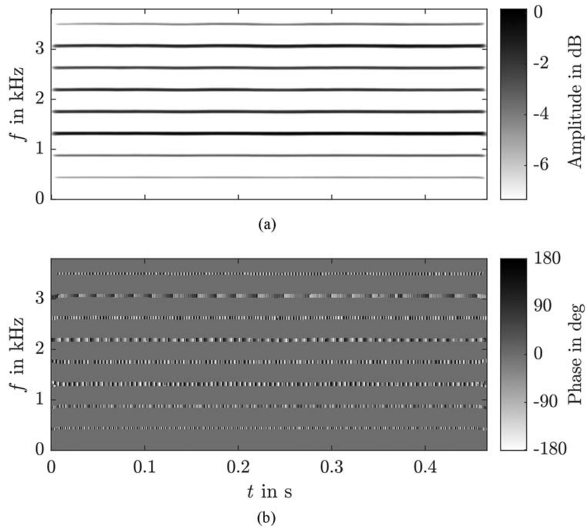

Fig. 2. Extract from the spectrogram for the signal of a trumpet playing note A4 ($f_0 = 440$ Hz) at fortissimo (microphone 4), with amplitude in decibels (a) and phase in degree (b).

center and small but certainly phase-relevant movements of the player. The non-stationary acoustic center is evident, for example, in deviations of the phase response in Fig. 2 from a strict periodicity that would be expected at constant travel time. Any reference, and thus any phase representation in the directivity of a musical instrument, would thus be arbitrary.

With regard to the use of directivity for room acoustic simulation and auralization, experience in physical rooms shows that small changes in the position of a sound source in space, although leading to substantial changes in the phase response of the TF, do not produce notable differences in sound, at least for natural, acoustic instruments. Thus, even for such applications, the phase can be chosen essentially freely.

Finally also the interpolation of the phase spectrum is highly susceptible to noise and can lead to errors, especially at high frequencies [19]. The authors have therefore proposed the use of absolute-valued directivities for interpolation [20]. This is in contrast to the previous processing of the data, which used the complex valued spectrum [8].

An estimate of the fundamental frequency f_0 was made by identifying the frequency with the highest amplitude within a window of ±100-cent bandwidth around the frequency corresponding to the tuning pitch indicated by the musicians. This was usually 442 or 443 Hz for modern instruments, 430 Hz for instruments of the Classical period, and 415 Hz for instruments of the Baroque period. This frequency was obtained for all 32 microphone recordings, and the most frequently occurring frequency over all 32 extracted values was chosen as f_0.

In the next step, the frequencies of the partials were estimated by placing a search window of ±10 cents around each harmonic frequency corresponding to f_0 and identifying the most frequently occurring, highest amplitude within this window, again considering all 32 microphone recordings. The search window takes into account the fact that, for physical reasons, the partials do not always lie exactly at the harmonic multiples of the fundamental frequency and that the fundamental was not always held exactly constant over the duration of the stationary part. If the identified frequency deviated from the harmonic frequencies by more than 5 cents, the procedure was stopped, and the signal energy was considered to be below the noise floor. A visual inspection of the detected signal components confirmed that all relevant and clearly identifiable partial tones had been found by this procedure.

At the frequencies f_i that could be estimated, the directivity was calculated using the power spectral density (PSD):

$$S_{xx,q}(k) = \frac{1}{f_s N} |X_q(k)|^2 , \tag{4}$$

which is a measure of the power of each frequency component in a signal with the implicit unit Pa^2/Hz. The authors have used the Welch method to improve the robustness of the PSD against noise by

1) Dividing the signal into eight segments of equal length with 50% overlap,
2) Applying a Hanning window function to each segment to reduce the effect of spectral leakage caused by the finite length of the segments,
3) Computing the periodogram of each segment, which is an estimate of the PSD of that segment, and
4) Averaging the periodograms across the segments to obtain an estimate of the overall PSD of the signal.

To obtain an estimate of the power at each frequency, each estimate of the PSD was scaled by the equivalent noise bandwidth of the window (in hertz). Finally the power was determined by simple peak picking in the scaled PSD around the frequencies f_i of the tone's harmonics and converted to the sound pressure value $p_q(f_i)$ by taking the square root. Fig. 3 illustrates the process for one note played by the modern oboe.

1.3.3 One-Third-Octave Band Averaging

Geometric acoustics applications typically require a single directivity pattern for a sound source, which is usually provided in one-third-octave band resolution. This was calculated by energetic averaging of the partials of all individual notes of an instrument falling into the one-third-octave bands according to IEC 61260-1 [21], for the $M = 30$ center frequencies from 25 Hz to 20 kHz. All partials of the J individual notes $p_q(f_{i,j})$ from Sec. 1.3.2 were used for the one-third-octave band representation. The data were averaged for each of the $Q = 32$ receivers of the spherical microphone array individually, by calculating the averaged amplitude as

$$\bar{p}_{q,m} = \sqrt{\frac{1}{L} \sum_{i=0}^{L-1} p_{q,i}^2}, \tag{5}$$

where L indicates the number of partials identified in one one-third-octave band. Fig. 4 illustrates the averaging pro-

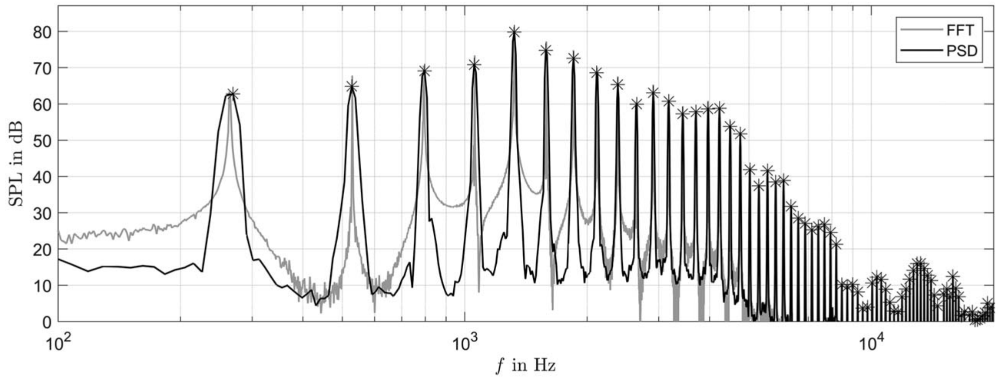

Fig. 3. Normalized FFT (gray) and scaled PSD (black) of the signal of the modern oboe playing note C4 (262 Hz) at fortissimo, recorded with microphone 4. The asterisks indicate the detected power values and frequencies f_0 and f_i that do not necessarily coincide with the discrete bins of the PSD. SPL = sound pressure level.

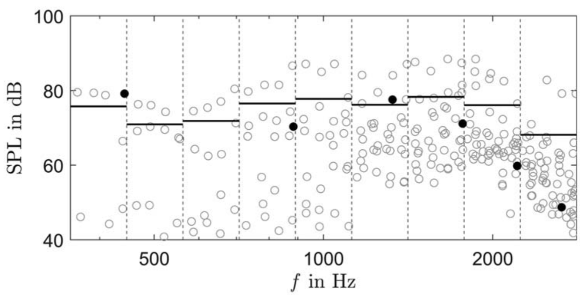

Fig. 4. Estimated sound pressure levels (SPLs) of all partials of the modern oboe plotted against frequency (gray) in the one-third-octave bands from 400 to 2.5 kHz, recorded with microphone 13. As an example, the fundamental and the first five overtones for the note A4 ($f_0 = 442$ Hz) are highlighted in black. The dashed vertical lines mark the boundaries of the one-third-octave bands. The horizontal solid black lines indicate the energetically averaged sound pressures $\bar{p}_{q,m}$ according to Eq. (5), shown in decibels.

cedure over all partials for the modern oboe in one-third-octave bands from 400 to 2.5 kHz.

At this point, the data still contain the direction-dependent frequency response of each instrument. If the directivities are used for auralizations in which a simulated (binaural) room impulse response is convolved with an anechoic recording of an instrument, it should be noted that the anechoic recording also contains the frequency response of the instrument in the direction of the recording microphone. To avoid an unnatural coloration during auralizations, the frequency response has to be removed from the one-third-octave band averaged directivities. In theory this normalization should also consider the position of the microphone from which the anechoic recording of the instrument was made, i.e., this direction should be normalized to 0 dB. If this position, however, is unknown or unstable due to movements of the instrument relative to the microphone, it may be most robust to equalize the directivity so that equal energy is radiated across all directions within each band, i.e., to

$$\bar{p}_{\mathrm{diff},q,m} = \frac{\bar{p}_{q,m}}{\sqrt{\sum_{q=1}^{Q} \bar{p}_{q,m}^2 \cdot w'_q}}, \tag{6}$$

where w'_q are the normalized area weights of the measurement grid with $\sum w'_q = 1$. This representation of the radiation patterns is called *diffuse* equalization in the following and will be discussed in more detail in SEC. 3.

The final step was to calibrate the directivity to the sound power of the real instruments, because the previous diffuse equalization [cf. Eq. (6)] had lost the absolute sound power reference. This was done by averaging the sound pressure level over the effective one-third-octave bands for each microphone. The one-third-octave bands with no sound radiation were set to zero.

The average sound pressure level of the diffuse equalized directivity per microphone was then calculated as

$$L_{\mathrm{P,3rd},q} = 10\lg\left(\frac{\frac{1}{M}\sum\limits_{m\in M}\bar{p}^2_{\mathrm{diff},q,m}}{p_0^2}\right), \tag{7}$$

where M indicates the number of effective one-third-octave bands and $p_0 = 2 \times 10^{-5}$ Pa. Its average over the surface of the spherical envelope is given by

$$\bar{L}_{\mathrm{P,3rd}} = 10\lg\left(\frac{1}{Q}\sum_{q=1}^{Q}10^{0.1L_{\mathrm{P,3rd},q}}\right). \tag{8}$$

The sound pressure of the reference, i.e., the corresponding instruments, was calculated from the calibrated recordings as follows:

$$L_{\mathrm{P,ref},q} = 10\lg\left(\frac{\frac{1}{NJ}\sum\limits_{n=1}^{N}\sum\limits_{j=1}^{J}x^2_{q,j}[n]}{p_0^2}\right), \tag{9}$$

where N is the number of samples of the stationary part and J is the number of single notes and averaged over all microphones to

$$\bar{L}_{\mathrm{P,ref}} = 10\lg\left(\frac{1}{Q}\sum_{q=1}^{Q}10^{0.1L_{\mathrm{P,ref},q}}\right). \tag{10}$$

Finally, the calibration of the diffuse equalized directivity from Eq. (6) is given by

$$\bar{p}_{\mathrm{cal},q,m} = \bar{p}_{\mathrm{diff},q,m} \cdot 10^{\frac{\bar{L}_{\mathrm{P,ref}} - \bar{L}_{\mathrm{P,3rd}}}{20}}. \tag{11}$$

An alternative way to calculate the one-third-octave band averaged directivity would have been to stitch all single-note recordings, use band filters, and analyze the complete signal within each one-third octave, which would result in an energy-weighted average of the individual frequency components. Due to the strong predominance of the tonal parts (see Fig. 3), this would, however, yield the same results compared to the authors' method within a range of precision that is reasonable for directivity measurements. The one-third-octave band averaged directivity of each instrument is calculated for the dynamic fortissimo (*ff*) and provided in the `FreeFieldDirectivityTF` SOFA convention.

1.3.4 Spatial Interpolation

Several applications that rely on the directivity of sound sources, such as room acoustic simulations, require the use of continuous or high resolution data. Consequently, the measurement data must be spatially resampled (interpolated) to match the required sampling grid.

By sampling the actual sound pressure function $f(\theta, \phi)$ with a Q channel spherical microphone array, the samples $p_q = f(\theta_q, \phi_q)$ are given at the positions (θ_q, ϕ_q) of the respective microphones for $q \in \{1, 2, ..., Q\}$. The general mathematical formula for interpolation can therefore be expressed as

$$\hat{f}(\theta_r, \phi_r) = \sum_{q=1}^{Q} f(\theta_q, \phi_q) \cdot L_q(\theta_r, \phi_r), \tag{12}$$

where $\hat{f}(\theta_r, \phi_r) = \hat{p}_r$ is the estimated sound pressure at the R points (θ_r, ϕ_r) of the interpolation grid for $r \in \{1, 2, ..., R\}$ and $L_q(\theta_r, \phi_r)$ is the interpolation function derived from the known sound pressure p_q at the position (θ_q, ϕ_q). The specific choice of the interpolation function depends on the interpolation method being used.

There is a plethora of techniques for interpolating real-valued scattered data that make different assumptions about the distribution of the discrete set of known data points [22]. For musical instruments, the thin-plate pseudo-spline method [23, 24] of order 1 has been found to be a good method, producing lower interpolation errors than spherical harmonics interpolation [25–27] or three-dimensional Vector-Based Amplitude Panning [28] when applied to sparsely sampled directivity measurements and evaluated against the directivity of different musical instruments measured at high resolution as a reference [20].

The authors chose an equiangular grid with an angular resolution of 5° in azimuth and colatitude as the target for the interpolation, resulting in $R = 2{,}522$ sound pressure values $\hat{p}_{r,m}$ for each of the $m \in M$ one-third-octave bands. The closed-form spherical spline interpolation for order 1 was realized with AKtools using the function `AKsphSplineInterp()` [29] based on the directivity patterns according to Eq. (11).

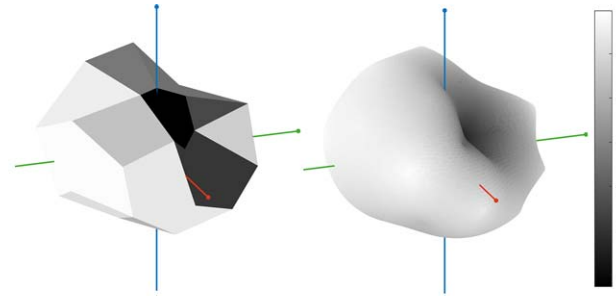

Fig. 5. The one-third-octave band averaged radiation pattern of the modern bassoon at 400 Hz, with original measurement resolution (32 points, a) and spatial interpolation to 2,522 points, using spherical splines of first order (right). Balloon representation with radius corresponding to the sound pressure level in decibels.

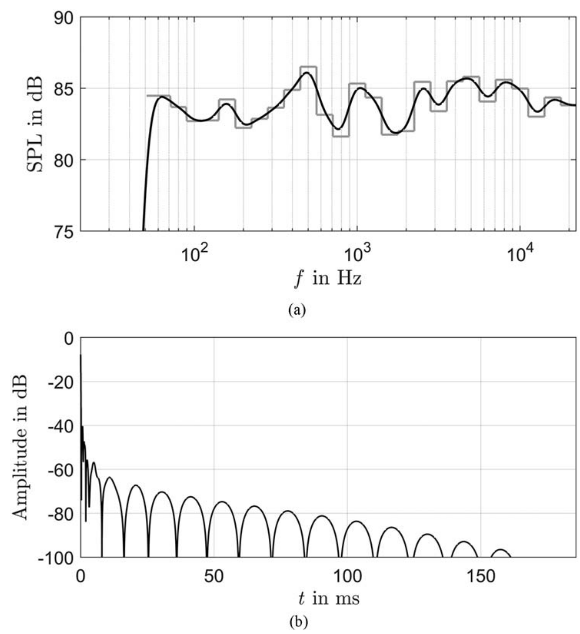

Fig. 6. The one-third-octave band spectrum (gray) and smoothed one-third-octave band spectrum (black) for the instrument bassoon modern (a) for the radiation direction ($\theta = 0$; $\phi = -10$) and corresponding energy-time curve (b). SPL = sound pressure level.

1.3.5 One-Third-Octave Band Band Smoothed FIRs

To enable musical instruments to be used as sound sources for simulating room acoustic and electroacoustic environments with software that uses FIR filters to represent directivity, $R = 2{,}522$ FIR filters were calculated using the above-mentioned grid. For this purpose, a one-third-octave band spectrum according to IEC 61260-1 [21] with a frequency resolution of 1 Hz was generated from the estimated sound pressure values $\hat{p}_{r,m}$ (gray line in Fig. 6). The one-sided spectrum with odd N was then smoothed with a one-third-octave band filter (black line in Fig. 6). After transformation to a two-sided spectrum according to Eq. (3), the spectrum was transformed into the time domain using inverse fast Fourier transform (IFFT), and the phase was made minimum phase using the AKTools function `AKphaseManipulation()`. Finally the FIR filter was reduced to 8,192 samples according to AES56-2008 [30] as shown in Fig. 5.

For use in the software EASE,[2] this data was stored in the proprietary GLL format.

2 DATABASE

The database contains the calibrated single-note recordings, single-note directivities, and frequency-averaged directivity patterns in SOFA format `FreeFieldDirectivityTF` convention under a Creative Commons share alike license (CC-BY-NC 4.0). In addition, high spatial resolution–interpolated radiation patterns averaged over one-third-octave bands are provided in openDaff format and as FIR filters in GLL format. All data are freely accessible [9]. A list of the available musical instruments can be found at Table 1.

The SOFA files can be read using a variety of APIs.[3] They contain the recorded signals and extracted directivities as complex TFs together with metadata describing the data in detail. This includes the name of the instrument in the entry `GLOBAL_SourceName`, the name of the musician in `GLOBAL_Musician`, the manufacturer of the instrument in `GLOBAL_SourceManufacturer`, and a verbal description of the position of the instrument during the measurement in `SourceView_Reference`. The arrangement of the capsules of the microphone array is described in `ReceiverPosition`.

In the SOFA data of the recordings, the respective note and dynamic level are indicated in `GLOBAL_Description`, e.g., `note = A4; dynamic =` *ff*, the MIDI number in `MidiNote`, e.g., 69 for A4, and the frequency for A4 corresponds to the tuning frequency in the entry for `SourceTuningFrequency`. In the SOFA data of the original recordings, `SteadyPart` indicates the range of the manually determined stationary portion in samples. A detailed description of the structure of the database can be found in the accompanying documentation.

2.1 SOFA Recordings

The single note recordings are available as a one-sided complex TF for each instrument and note. The data can be converted into a two-sided spectrum according to Eq. (3) and converted to the time domain by means of an IFFT. The naming of the data follows the scheme `SourceName_dynamic_midiNote_recordings.sofa`, and the recordings are stored in the `Data.Real` and `Data.Imag` fields. The fields have the dimension `MxRxN`, where M (measurement) is always 1; R (receiver) is the number of capsules of the microphone array, with R = 32; and N indicates the length of the TF. The field `N` contains the frequencies of the bins of the TF in hertz. Note that the calibrated 32-channel WAV recordings on which this data set is based are still freely accessible [9].[4]

2.2 SOFA Single Note

There is also a separate SOFA file for each instrument and note for the single-note directivity data; the naming of the file corresponds to the scheme `SourceName_midiNote_singleTones.sofa`. The purely real sound pressure levels are stored as complex TFs in the field `Data.Real` with the dimension `MxRxN`, where M is always 1, R = 32, and N refers to the number of the *I* extracted partials. The `Data.Imag` field with the dimension `MxRxN` is included in the dataset for consistency reasons but contains only zeros. The field `N` indicates the frequencies of the *I* partials in hertz.

2.3 SOFA One-Third-Octave Band

For the one-third-octave band averaged directivities, there is one SOFA file for each instrument with the naming scheme `SourceName_3rdOctave.sofa`. The averaged and calibrated sound pressures from Eq. (11) are stored in `Data.Real` with the dimension `MxRxN`, where M is always 1, R = 32, and N = 30 refers to the nominal center frequencies from 25 Hz to 20 kHz according to IEC 61260-1:2014 [21]. The data field `Data.Imag` has been filled with zeros. The field `N` indicates the center frequencies of the one-third-octave bands in hertz. In this case, the entries `SourceTuningFrequency`, `SteadyPart`, and `MidiNote` are not included in this data representation.

2.4 OpenDAFF One-Third-Octave Band

The open source format openDAFF can be read with several APIs.[5] The naming of the data follows the scheme `SourceName.daff`. The directional patterns are stored with a spatial resolution of 5° (azimuth and colatitude), i.e., each file contains one-third-octave band magnitude spectra at 2,522 points. These data can be used directly in the acoustic simulation environment RAVEN [31]. For evaluating the directivity, of both the individual notes and frequency-averaged directivity patterns, in arbitrary spatial resolution, the authors provide a MATLAB script as part of the database (cf. SEC. 2.6).

2.5 GLL One-Third-Octave Band FIRs

The proprietary GLL format allows the integration of complex sound sources into the acoustic simulation environment EASE. The directional patterns averaged over a one-third-octave band for all 21 modern musical instruments and for a soprano singer were stored as FIR filters with 8,192 taps and with a spatial resolution of 5°. For the exact naming scheme, please refer to the documentation of the dataset [9]. The data can be converted to UNF (used

---

[2] www.afmg.eu/en (accessed May 2, 2023).
[3] Cf. www.sofaconventions.org/mediawiki/index.php/Software_and_APIs (accessed May 2, 2023).
[4] In version 2 of the repository: https://doi.org/10.14279/depositonce-5861.2.
[5] Cf. www.github.com/svn2github/opendaff (accessed May 2, 2023).

by Ulysses[6]) and the CLF/CIF format (used by ODEON[7] and CATT-Acoustic[8]) using the proprietary SpeakerLab[9] software from AFMG.

2.6 Tools

Part of the database is the `Directivity_demo.m` MATLAB script that makes it possible to read the recordings from the SOFA data, display their spectrum graphically, and transform them into the time domain by IFFT. This data can then be saved as a WAV file and played back with common media players.

The script also allows for the three-dimensional display of single-note and frequency-averaged directivity in the form of balloon plots based on the SOFA data provided. Finally the data can be evaluated at any sampling quadrature using spherical spline interpolation. Prerequisite for all processing steps is the installation of AKTools[10] and the SOFA API for MATLAB (SOFAToolbox[11]) contained therein.

3 CONCLUSION

The present dataset contains recordings and radiation patterns of the individual notes of 41 modern and historical musical instruments, measured with a 32-channel microphone array in anechoic conditions. The recordings and directivities are provided in standardized SOFA format in the `FreeFieldDirectivityTF` convention. From these data, averaged directivities have been calculated for each instrument, which are suitable for use in acoustical simulation and auralization. In addition, spatially high-resolution directivities in OpenDAFF and GLL formats have been generated, allowing direct use in software such as RAVEN and EASE.

The absolute quality of the interpolation methods used for this spatial upsampling obviously depends on the characteristics of the sound source, such as its acoustically effective size, the modal patterns of its sound radiating parts, and the resulting complexity of the radiation pattern. For acoustically small sources, such as a trumpet and trombone bell or violin, a fairly accurate interpolation can be expected even based on a measurement at 32 points. For extended sources with more complex radiation patterns, however, a sparse sampling grid may lead to increasingly poor estimates of the far field directivity [20]. If other types of interpolation prove superior in the future, such as recently investigated, physically informed interpolation methods using the Euler equations as constraints [32], the interpolations applied may be revised in the future.

For the physically correct auralization of musical instruments in virtual acoustics, frequency-averaged directivities are a compromise that has to be made for technical reasons. Current simulation applications do not allow a straightforward exchange of single tone representations of such directivities within a simulation run. Due to the frequency averaging of the input data, a tonal coloration of the simulation result may occur [14]. The magnitude of this effect and its influence on instrumental and room acoustic perception will have to be determined in a subsequent study using the data presented.

When an anechoic recording of an instrument and its directivity is used to auralize a virtual acoustic environment, it is essential to normalize the directivity of the source to a suitable reference. In theory this would be the position of the microphone used to make the anechoic recording. Because the recorded signal already contains the directional timbre characteristic of the instrument in that direction, it should not be altered by the applied directivity, as this would result in unwanted coloration of the simulation. This means that the directivity should be normalized, so that only a frequency response *relative* to this reference direction will be obtained. This will be achieved by calculating

$$\hat{p}_{\mathrm{pt},r,m} = \frac{\hat{p}_{r,m}}{\hat{p}_{\mathrm{mic},m}}, \tag{13}$$

where $\hat{p}_{\mathrm{mic},m}$ is the interpolated and frequency-averaged sound pressure of the M one-third-octave band in the reference direction.

In a real recording situation, however, the instrument being played by a musician is a moving sound source. This can result in an angular displacement of the instrument which can easily reach up to 47° when played in a standing position and up to 36° when played in a sitting position [33]. To compensate for the movement of the instrument when referencing, the authors consider it useful to equalize the directivity not to a point but to the average of a larger spherical surface area A. The distribution of orientations over this surface can be taken into account by using a weighting function, i.e., by calculating

$$\hat{p}_{\mathrm{area},r,m} = \frac{\hat{p}_{r,m}}{\sqrt{\sum_{r_A=1}^{R_A} \hat{p}^2_{r_A,m} \cdot w'_{r_A} \cdot g'_{r_A}}}, \tag{14}$$

where w'_{r_A} and g'_{r_A} are the normalized area and two-dimensional function weights, respectively, with

$$\sum_{r_A=1}^{R_A} w'_{r_A} \cdot g'_{r_A} = 1, \tag{15}$$

and $r_A \in \{1, 2, ..., R_A\}$ are the R_A grid points of the area A over which the mean is to be calculated.

If no information about the position of the microphone during the recording is documented, a diffuse equalized directivity [cf. Eq. (6)] can be used to minimize the sound coloration on average. For this reason, the directivity averaged over one-third-octave bands has been provided in this representation for uniformed immediate use. The extent to

[6] www.ifbsoft.de (accessed May 2, 2023).
[7] www.odeon.dk (accessed May 2, 2023).
[8] www.catt.se (accessed May 2, 2023).
[9] www.afmg.eu/en/ease-speakerlab (accessed May 2, 2023).
[10] Cf. www.tu.berlin/ak/forschung/publikationen/open-research-tools/aktools (accessed May 2, 2023).
[11] Cf. www.github.com/sofacoustics/SOFAtoolbox (accessed May 02, 2023).

which the different referencing methods of the recording position affect the perceived sound event and room acoustic parameters will be clarified in a later investigation on the basis of these data.

4 ACKNOWLEDGMENT

This work was funded by the German Research Foundation (DFG WE 4057/9-2). The authors would like to thank Stefan Feistel and Silke Bögelein (AFMG) for their support in producing the GLL datasets.

5 REFERENCES

[1] H. K. Dunn and D. W. Farnsworth, "Exploration of Pressure Field Around the Human Head During Speech," *J. Acoust. Soc. Am.*, vol. 10, no. 3, pp. 184–199 (1939 Jan.). https://doi.org/10.1121/1.1915975.

[2] J. Meyer, *Acoustics and the Performance of Music* (Springer, New York, 2009).

[3] O. Abe, *Sound Radiation of Singing Voices*, Ph.D. thesis, Universität Hamburg, Hamburg, Germany (2019 Sep.).

[4] C. Pörschmann, "A Database for the Comparison of Measured Datasets of Human Voice Directivity," in *Proceedings of the Forum Acusticum* , pp. 1–8 (Torino, Italy) (2023 Sep.).

[5] F. Hohl, *Kugelmikrofonarray zur Abstrahlungsver-Messung von Musikinstrumenten*, Master's thesis, Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz, Austria (2009 Nov.).

[6] J. Pätynen and T. Lokki, "Directivities of Symphony Orchestra Instruments," *Acta Acust. united Acust.*, vol. 96, no. 1, pp. 138–167 (2010 Jan./Feb.). https://doi.org/10.3813/AAA.918265.

[7] S. D. Bellows, K. J. Bodon, and T. W. Leishman, "Musical Instrument Directivity," https://scholarsarchive.byu.edu/directivity/ (accessed Jul. 20, 2020).

[8] N. R. Shabtai, G. Behler, M. Vorländer, and S. Weinzierl, "Generation and Analysis of an Acoustic Radiation Pattern Database for Forty-One Musical Instruments," *J. Acoust. Soc. Am.*, vol. 141, no. 2, pp. 1246–1256 (2017 Feb.). https://doi.org/10.1121/1.4976071.

[9] S. Weinzierl, M. Vorländer, D. Ackermann, et al., "A Database of Anechoic Microphone Array Measurements of Musical Instruments," *DepositOnce: Technische Universität Berlin, 5861, Version 3* (2024 Feb.). https://doi.org/10.14279/depositonce-19858.

[10] AES, "AES Standard for File Exchange - Spatial Acoustic Data File Format," *AES Standard 69-2022* (2022 Sep.).

[11] P. Majdak, F. Zotter, F. Brinkmann, et al., "Spatially Oriented Format for Acoustics 2.1: Introduction and Recent Advances," *J. Audio Eng. Soc.*, vol. 70, no. 7, pp. 565–584 (2022 Jul.). https://doi.org/10.17743/jaes.2022.0026.

[12] F. Wefers, "OpenDAFF: A Free, Open-Source Software Package for Directional Audio Data," in *Proceedings of the Fortschritte der Akustik*, pp. 1059–1060 (Berlin, Germany) (2010 Mar.).

[13] AFMG, "GLL (Generic Loudspeaker Library) – A New Standard for Measuring and Storing Loudspeaker Performance Data," *AFMG Engineering White Paper* (2007 Oct.). https://www.afmg.eu/sites/default/files/2021-07/GLL_White_Paper_October07.pdf.

[14] A. Corcuera, V. Chatziioannou, and J. Ahrens, "Perceptual Significance of Tone-Dependent Directivity Patterns of Musical Instruments," *J. Audio Eng. Soc.*, vol. 71, no. 5, pp. 293–302 (2023 May). https://doi.org/10.17743/jaes.2022.0076.

[15] M. Kob, "Impact of Excitation and Acoustic Conditions on the Accuracy of Directivity Measurements," *Proceedings of theInternational Symposium on Musical Acoustics (ISMA)*, pp. 639–643 (Le Mons, France) (2014 Jul.).

[16] S. Weinzierl, S. Lepa, F. Schultz, et al., "Sound Power and Timbre as Cues for the Dynamic Strength of Orchestral Instruments," *J. Acoust. Soc. Am.*, vol. 144, no. 3, pp. 1347–1355 (2018 Sep.). https://doi.org/10.1121/1.5053113.

[17] J. Ahrens, C. Andersson, P. Höstmad, and W. Kropp, "Tutorial on Scaling of the Discrete Fourier Transform and the Implied Physical Units of the Spectra of Time-Discrete Signals," presented at the *148th Convention of the Audio Engineering Society* (2020 May), e–Brief 600.

[18] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The Timbre Toolbox: Extracting Audio Descriptors From Musical Signals," *J. Audio Eng. Soc.*, vol. 130, no. 5, pp. 2902–2916 (2011 Nov.).

[19] J. Ahrens and S. Bilbao, "Computation of Spherical Harmonics Based Sound Source Directivity Models From Sparse Measurement Data," in *Proceedings of the Forum Acusticum*, pp. 2019–2026 (Lyon, France) (2020 Dec.).

[20] D. Ackermann, F. Brinkmann, F. Zotter, M. Kob, and S. Weinzierl, "Comparative Evaluation of Interpolation Methods for the Directivity of Musical Instruments," *EURASIP J. Audio Speech Music Process.*, vol. 2021, paper 36 (2021 Oct.). https://doi.org/10.1186/s13636-021-00223-6.

[21] IEC, "Electroacoustics - Octave-Band and Fractional-Octave-Band Filters - Part 1: Specifications," *IEC Standard 61260-1:2014* (2014 Jun.).

[22] J. Li and A. D. Heap, *A Review of Spatial Interpolation Methods for Environmental Scientists* (Geoscience, Australia, Canberra, 2008).

[23] G. Wahba, "Spline Interpolation and Smoothing on the Sphere," *SIAM J. Sci. Stat. Comput.*, vol. 2, no. 1, pp. 5–16 (1981 Mar.). https://doi.org/10.1137/0902002.

[24] G. Wahba, "Erratum: Spline Interpolation and Smoothing on the Sphere," *SIAM J. Sci. Stat. Comput.*, vol. 3, no. 3, pp. 385–386 (1982 Sep.). https://doi.org/10.1137/0903024.

[25] G. Weinreich and E. B. Arnold, "Method for Measuring Acoustic Radiation Fields," *J. Acoust. Soc. Am.*, vol. 68, no. 2, pp. 404–411 (1980 Aug.). https://doi.org/10.1121/1.384751.

[26] F. Zotter, *Analysis and Synthesis of Sound-Radiation With Spherical Arrays*, Ph.D. thesis, University of Music and Performing Arts, Graz, Austria (2009 Nov.).

[27] M. Pollow, *Directivity Patterns for Room Acoustical Measurements and Simulations*, Ph.D. thesis, RWTH Aachen University, Aachen, Germany (2014 Sep.).

[28] V. Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466 (1997 Jun.).

[29] F. Brinkmann and S. Weinzierl, "AKtools — An Open Software Toolbox for Signal Acquisition, Processing, and Inspection in Acoustics," presented at the *142nd Convention of the Audio Engineering Society* (2017 May), eBrief 309.

[30] AES, "AES Standard on Acoustics - Sound Source Modeling - Loudspeaker Polar Radiation Measurements," *AES Standard 56-2008* (2019 Aug.).

[31] D. Schröder and M. Vorlaender, "RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environments," in *Proceedings of the Forum Acusticum*, pp.1541–1546 (Aalborg, Denmark) (2011 Jun./Jul.).

[32] M. Lemke, A. Hoelter, D. Ackermann, and S. Weinzierl, "Physics-Informed Interpolation of Directional Characteristics," presented at the *Forum Acusticum* (Torino, Italy) (2023 Sep.).

[33] D. Ackermann, C. Böhm, F. Brinkmann, and S. Weinzierl, "The Acoustical Effect of Musicians' Movements During Musical Performances," *Acta Acust. united Acust.*, vol. 105, no. 2, pp. 356–367 (2019 Mar./Apr.). https://doi.org/10.3813/AAA.919319.

THE AUTHORS

David Ackermann

Fabian Brinkmann

Stefan Weinzierl

David Ackermann received his M.Sc. degree in Audiocommunication and Technology from Technische Universität Berlin, Germany, in 2015. Since then he has been a research associate in the Audio Communication Group at the Technische Universität Berlin, where his research is focused in the field of virtual and musical acoustics. His research includes the investigation of the time-dependent behavior of natural acoustic sound sources and their auralization in virtual acoustic environments.

•

Fabian Brinkmann received an M.A. degree in Communication Sciences and Technical Acoustics in 2011 and Dr. rer. nat. degree in 2019 from the Technische Universität Berlin, Germany. He focuses on the fields of signal processing and evaluation approaches for spatial audio. Fabian is a member of the AES, German Acoustical Society (DEGA), and European Acoustics Association (EAA) technical committee for psychological and physiological acoustics.

•

Stefan Weinzierl is head of the Audio Communication Group at the Technische Universität Berlin. His research is focused on audio technology, virtual acoustics, room acoustics, and musical acoustics. With a diploma in physics and sound engineering, he received his Ph.D. in musical acoustics. He is coordinating a master program in Audio Communication and Technology at Technische Universität Berlin and has coordinated international research consortia in the field of virtual acoustics (SEACEN) and music information retrieval (ABC_DJ).