Aliasing-Free Time-Domain Simulation of Spherical Microphone Arrays

Nara Hahn[1,2], Frank Schultz[1], and Sascha Spors[1]

[1] *Institute of Communications Engineering, University of Rostock, Rostock, Germany*

[2] *Institute of Sound and Vibration Research, University of Southampton, Southampton, United Kingdom*

Email: nara.hahn@soton.ac.uk, {frank.schultz, sascha.spors}@uni-rostock.de

Introduction

This paper proposes an improved time-domain simulation of rigid spherical microphone arrays. The spatial transfer function of an external sound source evaluated on a rigid sphere can be represented as a spherical harmonics expansion. Each harmonic mode is described by the spherical Hankel function and its derivative, where Laplace-domain poles and zeros characterize the corresponding spectrum [1–4]. We convert the continuous-time transfer function into the z-domain using a recently proposed discretization technique called the band-limited impulse invariance method [5]. The resulting discrete-time system has a parallel structure composed of digital IIR filters and a single FIR filter. The FIR part is designed to cancel the aliasing introduced by the IIR filters. The captured signals are simulated by applying the filters to any given source signal. The filter coefficients are given in closed form, which is derived analytically. The design accuracy of the proposed method is demonstrated by numerical simulations. The results are compared with frequency-domain approach and the conventional impulse invariance method.

Time-Domain Simulation

The sound field of a point source at $\boldsymbol{x}_\mathrm{s} = (r_\mathrm{s}, \theta_\mathrm{s}, \phi_\mathrm{s})$ captured on a rigid sphere $\boldsymbol{x} = (R, \theta, \phi)$ at the origin is expressed in the spherical harmonics domain as ($R < r_\mathrm{s}$) [6, Sec. 4.2]

$$S(\boldsymbol{x}, \omega) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} P_n\left(\cos\Theta_\mathrm{s}\right) \frac{-h_n(\frac{\omega}{c} r_\mathrm{s})}{\frac{\omega}{c} R^2 \cdot h_n'(\frac{\omega}{c} R)}. \quad (1)$$

The spherical coordinates (r, θ, ϕ) are defined with colatitude $\theta \in [0, \pi]$ and azimuth $\phi \in [0, 2\pi)$. The axis-symmetric part in (1), described by the Legendre polynomial $P_n(\cdot)$, is dependent on the angle between the source and the receiver denoted by Θ_s. The spherical Hankel functions of the second kind $h_n(\cdot)$ and the derivatives $h_n'(\cdot)$ with respect to its argument characterize the distance- and frequency-dependent part of the individual modes. The angular frequency ω in rad/s is related to frequency in Hz by $\omega = 2\pi f$. The imaginary unit is denoted by i and the speed of sound in m/s by c.

The system function of each mode can be derived by exploiting the series expansion of the spherical Hankel functions $h_n(\cdot)$,

$$h_n(\xi) = -\frac{i^n e^{-i\xi}}{(i\xi)^{n+1}} \sum_{k=0}^{n} \beta_n(k) \cdot (i\xi)^k, \quad (2)$$

where the coefficients are given as

$$\beta_n(k) = \begin{cases} \frac{(2n-k)!}{(n-k)!\ k!\ 2^{n-k}}, & k = 0, 1, \ldots, n \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The expression (2) can be obtained by formulating [7, Eq. (10.49.7)] in terms of $i\xi$. Similarly, the derivative of the spherical Hankel function $h_n'(\cdot)$ can be expressed as a series expansion by using the recurrence relation [7, Eq. (10.51.2)],

$$h_n'(\xi) = -h_{n+1}(\xi) + \frac{n}{\xi} h_n(\xi), \quad (4)$$

yielding

$$h_n'(\xi) = \frac{i^{n+1} e^{-i\xi}}{(i\xi)^{n+2}} \sum_{k=0}^{n+1} \gamma_n(k) \cdot (i\xi)^k, \quad (5)$$

where the expansion coefficients are

$$\gamma_n(k) = \begin{cases} \beta_{n+1}(k) - n \cdot \beta_n(k), & k = 0, 1, \ldots, n \\ 1, & k = n+1 \\ 0, & \text{otherwise} \end{cases}. \quad (6)$$

By using (2) and (5), the rational part in (1) can be rewritten as

$$\frac{-h_n(\frac{\omega}{c} r_\mathrm{s})}{\frac{\omega}{c} R^2 \cdot h_n'(\frac{\omega}{c} R)} = \frac{c}{r_\mathrm{s} R} e^{-i\frac{\omega}{c}(R - r_\mathrm{s})} A_n(i\omega), \quad (7)$$

which consists of an amplitude scaling $\frac{c}{r_\mathrm{s} R}$, a propagation delay of $\frac{r_\mathrm{s} - R}{c}$ seconds, and a frequency-dependent term denoted by $A_n(i\omega)$. The modal spectrum $A_n(i\omega)$ is characterized by the ratio of two polynomials,

$$A_n(i\omega) = \frac{\sum_{k=0}^{n} \hat{\beta}_n(k) \cdot (i\omega)^k}{\sum_{k=0}^{n+1} \hat{\gamma}_n(k) \cdot (i\omega)^k}, \quad (8)$$

whose coefficients are defined as

$$\hat{\beta}_n(k) = \beta_n(k) \cdot (\tfrac{r_\mathrm{s}}{c})^{k-n} \quad (9)$$

$$\hat{\gamma}_n(k) = \gamma_n(k) \cdot (\tfrac{R}{c})^{k-n-1}. \quad (10)$$

The magnitude spectra for some selected orders ($n = 0, 1, \ldots, 7$) are shown in Fig. 3 (top left). The 0th-order spectrum constitutes a first-order low-pass filter. For $n \geq 1$, the modal spectrum has its peak around $f \approx \frac{c\,n}{2\pi R}$. The magnitude response has a low-shelf cut at lower frequencies, and exhibits a low-pass slope of -20 dB/decade at higher frequencies.

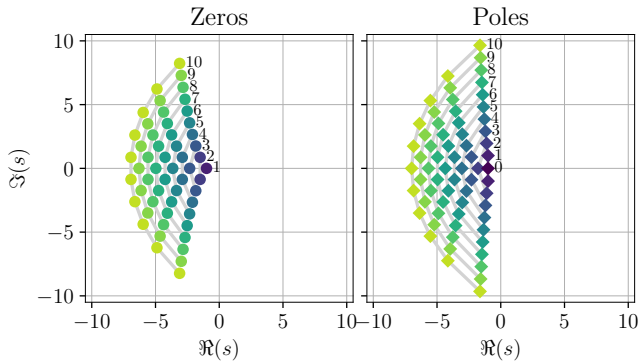

Figure 1: Laplace-domain zeros and poles of the modal transfer functions for $n = 0, 1, \ldots, 10$ and $\frac{r_\mathrm{s}}{c} = \frac{R}{c} = 1\,\mathrm{s}$ (cf. (11)). The zeros/poles of each order are depicted with the same color and connected with gray lines. The digits indicate the respective modal orders. Note that, for $n = 0$, the modal transfer function is a single-pole system with no zero.

The analytic continuation of (8) can be considered as the Laplace-domain representation ($s \in \mathbb{C}$) of the system function,

$$A_n(s) = \frac{\sum_{k=0}^{n} \hat{\beta}_n(k)\, s^k}{\sum_{k=0}^{n+1} \hat{\gamma}_n(k)\, s^k}, \tag{11}$$

meaning that $A_n(i\omega)$ is the slice of $A_n(s)$ along the imaginary axis ($s = i\omega$). The nth-order modal spectrum is therefore characterized by n zeros and $n+1$ poles, which corresponds to the roots of the numerator and the denominator, respectively. The zeros and poles of the first eleven modal spectra ($n = 0, \ldots, 10$) are depicted in Fig. 1 for $\frac{r_\mathrm{s}}{c} = \frac{R}{c} = 1\,\mathrm{s}$ (and thus $\hat{\beta}_n(k) = \beta_n(k)$ and $\hat{\gamma}_n(k) = \gamma_n(k)$). It can be seen that the roots are distinct and lie on the left half-plane ($\Re(s) < 0$). The modal transfer functions are thus causal and minimum-phase [8].

In the current study, the modal transfer function (11) is converted into a partial fraction expansion,

$$A_n(s) = \sum_{k=0}^{n+1} \frac{\rho_k}{s - p_k}, \tag{12}$$

where p_k denotes the poles and ρ_k the residues. Equation (12) constitutes a parallel-structure filter. For a digital emulation of the continuous-time transfer function $A_n(s)$, the Laplace-domain representation has to be converted into the z-domain. Conventionally used conversion methods are the bilinear transform, the impulse invariance method, and the matched z-transform. The accuracy of the resulting digital filters are limited, as demonstrated in [4]. An improved digital modeling is introduced in the following section.

Band-Limited Impulse Invariance Method

The band-limited impulse invariance method was proposed in [5], which models a continuous-time system with high accuracy. It is an extension of the conventional impulse invariance method which is one of the standard mappings from the Laplace domain to the z-domain [9]. It is based on time-domain sampling of continuous-time impulse responses. In the presented method, the modeling accuracy in terms of magnitude and phase spectra is improved by applying an analytical anti-aliasing filtering prior to the time-domain sampling. This section briefly introduces the band-limited impulse invariance method, where a first-order system with a single pole is considered. Applying it to higher-order systems (e.g. (12)) is straightforward. Interested readers are referred to [5] for further details.

We assume that the first-order system is causal and expressed in the Laplace-domain as

$$H_\mathrm{a}(s) = \frac{\rho}{s - p}, \tag{13}$$

where the pole and residue are respectively noted by p and ρ ($\Re(p) < 0$). For a sampling frequency f_s, an ideal anti-aliasing filtering yields a low-pass filtered system function,

$$H_\mathrm{BL}(s) = \begin{cases} \frac{\rho}{s-p}, & |\Im(s)| < \frac{\omega_\mathrm{s}}{2} \\ 0, & |\Im(s)| > \frac{\omega_\mathrm{s}}{2} \end{cases}, \tag{14}$$

with $\Im(\cdot)$ denoting the imaginary part of a complex variable and $\frac{\omega_\mathrm{s}}{2} = \pi f_\mathrm{s}$ the Nyquist angular frequency in rad/s.

The ideally low-pass filtered impulse response can then be obtained by deriving the inverse Laplace transform of (14), yielding

$$h_\mathrm{BL}(t) = \frac{1}{2\pi i} \int_{-i\frac{\omega_\mathrm{s}}{2}}^{+i\frac{\omega_\mathrm{s}}{2}} \frac{\rho}{s - p} \mathrm{d}s \tag{15}$$

$$= \rho \cdot e^{pt} u(t) + \rho \cdot \varepsilon(t), \tag{16}$$

where $u(t)$ denotes the Heaviside step function. The residual function, denoted by $\varepsilon(t)$, is defined as

$$\varepsilon(t) \tag{17}$$

$$= \begin{cases} \frac{e^{pt}}{2\pi i}\left[E_1\left((i\frac{\omega_\mathrm{s}}{2} + p)t\right) - E_1\left((-i\frac{\omega_\mathrm{s}}{2} + p)t\right)\right], & t \neq 0 \\ \frac{1}{2\pi i}\left[\log(i\frac{\omega_\mathrm{s}}{2} - p) - \log(-i\frac{\omega_\mathrm{s}}{2} - p)\right] - u(0), & t = 0 \end{cases}$$

where $E_1(z) := \int_z^\infty \frac{e^{-z'}}{z'} \mathrm{d}z'$ is the exponential integral function [7, Eq. (6.2.1)]. Note that the decaying exponential term $\rho \cdot e^{pt} u(t)$ in (16) corresponds to the impulse response of (13) without low-pass filtering, i.e. the full-band impulse response. Equation (16) suggests that an anti-aliasing filtering can be carried out by superimposing the residual function $\rho \cdot \varepsilon(t)$ onto the full-band impulse response. This also means that the residual function constitutes an ideally high-pass filtered impulse response with reversed sign (out of phase),

$$\mathcal{E}(s) = \begin{cases} 0, & |\Im(s)| < \frac{\omega_\mathrm{s}}{2} \\ -\frac{1}{s-p}, & |\Im(s)| > \frac{\omega_\mathrm{s}}{2} \end{cases}, \tag{18}$$

where $\mathcal{E}(s)$ denotes the Laplace transform of $\varepsilon(t)$. When $\rho \cdot \mathcal{E}(s)$ is added to the original transfer function (13), the

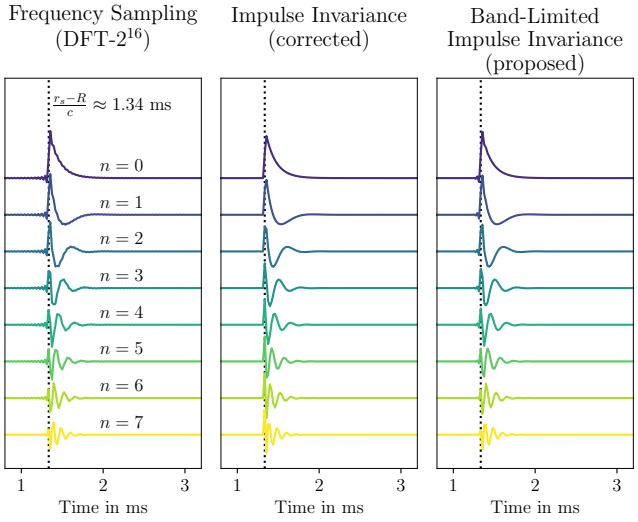

Figure 2: Modal impulse responses obtained from different simulation approaches ($R = 0.042\,\mathrm{m}$, $r_\mathrm{s} = 0.5\,\mathrm{m}$, $n = 0, \ldots, 7$, $f_\mathrm{s} = 48\,\mathrm{kHz}$). The dotted line indicates the time-of arrival of the wavefront on the sphere ($t_\mathrm{TOA} = \frac{r_\mathrm{s}-R}{c} \approx 1.34\,\mathrm{ms}$). The curves are vertically separated for the ease of visualization.

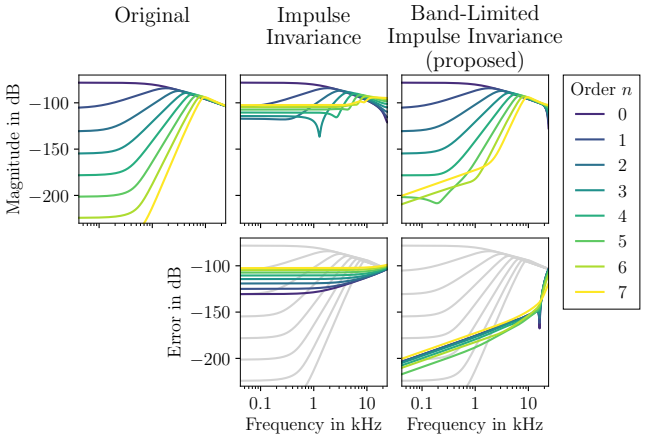

Figure 3: Modal magnitude spectra obtained from different simulation approaches ($R = 0.042\,\mathrm{m}$, $r_\mathrm{s} = 0.5\,\mathrm{m}$, $n = 0, \ldots, 7$, $f_\mathrm{s} = 48\,\mathrm{kHz}$) and the resulting spectral errors. The top-left plot depicts the ideal modal spectra, which are also shown in the bottom row (gray).

spectrum above the Nyquist limit ($|\omega| > \frac{\omega_\mathrm{s}}{2}$) is canceled, yielding the ideally low-pass filtered system function as in (14).

When sampled with an interval of $T_\mathrm{s} := \frac{1}{f_\mathrm{s}}$, the full-band impulse response and the residual exhibit aliasing spectra with the same magnitude spectrum but out of phase. Thus, superimposing the respective signals in the discrete-time domain cancels the aliasing. The full-band impulse response is realized by an IIR filter using the conventional impulse invariance method [10, 11]. The z-transform of the resulting filter reads

$$H^{(\mathrm{IIR})}(z) = \frac{\rho T_\mathrm{s}}{2} \frac{1 + e^{pT_\mathrm{s}} z^{-1}}{1 - e^{pT_\mathrm{s}} z^{-1}}, \tag{19}$$

which exhibits a pole at $z = e^{pT_\mathrm{s}}$ and a zero at $z = -e^{pT_\mathrm{s}}$. The residual is realized by an FIR filter whose coefficients are obtained by sampling $\rho \cdot \varepsilon(t)$. To reduce the length of the noncausal part, the residual has to be truncated and windowed to a finite length (L samples) before the sampling.

The first-order system (13) is modeled by a parallel combination of the IIR and FIR filter. The IIR filter output has to be delayed to accommodate the noncausal part ($M \leq L$) of the FIR filter. The z-domain transfer function of the digital filter thus reads

$$H_\mathrm{d}(z) = H^{(\mathrm{FIR})}(z) + z^{-M} H^{(\mathrm{IIR})}(z). \tag{20}$$

The length of the FIR filter L and the overall delay M are design parameters, which trades off the computational complexity of the accuracy of the discrete-time model.

Evaluation

The proposed method is evaluated for a rigid sphere with radius $R = 0.042\,\mathrm{m}$. We consider the sound field of a spherical wave emitted by a point source at $\boldsymbol{x}_\mathrm{s} = (0.5\,\mathrm{m}, 90°, 180°)$, i.e. on the negative x-axis. The speed of sound is assumed to be $c = 343\,\mathrm{m/s}$, and the sampling frequency is set to $f_\mathrm{s} = 48\,\mathrm{kHz}$.

For the band-limited impulse invariance method, the FIR length is set to 15 samples and the pre-delay to 7 samples. The FIR part is designed by windowing the residue with a Kaiser-Bessel window ($\beta = 8.6$). The result is compared with those from the frequency-domain sampling and the conventional impulse invariance method. For the frequency-domain simulation, the DFT length is set to 2^{16}. The corrected version of the impulse invariance method is used as described in [12, Sec. 7.3] and [10, 11]. Note that each method results in a different group delay. In the following evaluations, the group delays are removed and the impulse responses are time aligned for ease of comparison.

The point source is excited by a discrete-time impulse $\delta[n]$. The resulting modal impulse responses and modal spectra for the first eight orders ($n = 0, 1, \ldots, 7$) are depicted in Fig. 2 and Fig. 3, respectively. The errors shown in Fig. 3 (bottom row) are the spectral deviations from the original modal spectra (top left). Lower error thus means higher accuracy in magnitude and phase spectrum.

In Fig. 2, the time-of-arrival of the first wavefront on the spherical surface is indicated by dotted lines ($t_\mathrm{TOA} = \frac{r_\mathrm{s}-R}{c} \approx 1.34\,\mathrm{ms}$). Prominent pre- and post-ringings are observed for the frequency-domain design (first column) which is mainly due to the brick-wall low-pass filtering of the spectrum below the Nyquist frequency $\frac{f_\mathrm{s}}{2}$. Note also that noncausal components appear earlier than t_TOA, leading to a longer group delay than other methods. The conventional impulse invariance method (second column), which is based on time-domain sampling, does not exhibit such oscillation and yields strictly causal impulse responses. This causality, however, comes at the expense of spectral errors as shown in Fig. 3. The proposed

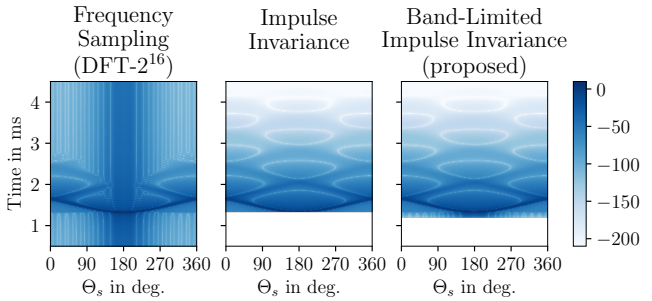

Figure 4: Spatial impulse responses on the rigid sphere ($R = 0.042\,\mathrm{m}$, $r_\mathrm{s} = 0.5\,\mathrm{m}$, $n = 0, \ldots, 18$, $f_\mathrm{s} = 48\,\mathrm{kHz}$). The receivers are placed on the equator, i.e. $\theta = 90°$, $\phi \in [0°, 360°)$.

method (last column) is a compromise between these two extreme cases. The FIR part reduces the aliasing occurring in the impulse invariance method thereby improving the modeling accuracy (cf. Fig. 3). Further improvements can be achieved by increasing the FIR length L (not shown here). The length of the noncausal part can be controlled explicitly by the design parameter M.

The spatial impulse responses on the sphere are computed by evaluating the modal expansion up to the 18th order. To alleviate the truncation effect, a modal smoothing is applied with the right-half of the Kaiser-Bessel window ($\beta = 8.6$). Fig. 4 depicts the responses on the equator ($\theta = 90°$, $\phi \in [0°, 360°)$). It can be seen that the frequency-domain simulation exhibits strong temporal artifacts due to the brick-wall low-pass filtering and temporal aliasing. Both the conventional and band-limited impulse invariance methods better reveal the spatio-temporal structure of the sound field. The diffraction pattern of the wavefront is clearly visible including the creeping waves which propagate along the surface [13, 14]. The two impulse invariance methods differ only within the range $t \in [t_\mathrm{TOA} - M \cdot T_\mathrm{s}, t_\mathrm{TOA} + (L - M) \cdot T_\mathrm{s}]$ where FIR filters are used by the band-limited impulse invariance method. Note that the noncausal component ($t \in [t_\mathrm{TOA} - M \cdot T_\mathrm{s}, t_\mathrm{TOA}]$) resulting from the band-limited impulse invariance method resembles that of the frequency-domain method. This is not surprising, as those components are essential for an aliasing-free simulation.

Conclusion

This paper proposes a time-domain simulation of sound fields captured on a rigid sphere. A recently introduced discretization technique, called the band-limited impulse invariance method, is used. The modeling accuracy in the frequency domain is improved by reducing the frequency-domain aliasing which typically occurs during time-domain sampling. It is also demonstrated that the temporal structure of the sound field is better modeled compared to the frequency-sampling method. The improvement is achieved at the cost of additional FIR filters whose length can be freely chosen by the user depending on the desired accuracy, available computational power, and allowable latency (group delay). This approach might be used for an efficient simulation of spherical microphone arrays which are frequently used in sound field capturing.

References

[1] H. Pomberger, "Angular and radial directivity control for spherical loudspeaker arrays," Master's thesis, University of Music and Performing Arts, Graz, Austria, 2008.

[2] F. Zotter, "Analysis and synthesis of sound-radiation with spherical arrays," Ph.D. dissertation, University of Music and Performing Arts, Graz, Austria, 2009.

[3] S. Lösler and F. Zotter, "Comprehensive radial filter design for practical higher-order Ambisonic recording," in *Proc. 41st German Annu. Conf. Acoust. (DAGA)*, 2015, pp. 452–455.

[4] N. Hahn, F. Schultz, and S. Spors, "Spatio-temporal properties of simulated impulse responses on a rigid sphere," in *Proc. 47th German Annu. Conf. Acoust. (DAGA)*, Vienna, Austria, Aug. 2021.

[5] ——, "Band limited impulse invariance method," in *Proc. 30th Eur. Signal Process. Conf. (EUSIPCO)*, Belgrade, Serbia, Sep. 2022, pp. 209–213.

[6] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Oxford, UK: Elsevier, 2005.

[7] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions Handback*. New York, NY, USA: Cambridge University Press, 2010.

[8] B. Girod, R. Rabenstein, and A. Stenger, *Signals and Systems*. Wiley, 2001.

[9] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Prentice Hall, 1999.

[10] L. B. Jackson, "A correction to impulse invariance," *IEEE Signal Process. Lett.*, vol. 7, no. 10, pp. 273–275, 2000.

[11] W. F. Mecklenbräuker, "Remarks on and correction to the impulse invariant method for the design of IIR digital filters," *Signal Process.*, vol. 80, no. 8, pp. 1687–1690, 2000.

[12] R. A. Gabel and R. A. Roberts, *Signals and Linear Systems*. New York, NY, USA: Wiley, 1973.

[13] F. M. Wiener, "Sound diffraction by rigid spheres and circular cylinders," *J. Acoust. Soc. Am. (JASA)*, vol. 19, no. 3, pp. 444–451, 1947.

[14] M. Harbold and B. Steinberg, "Direct experimental verification of creeping waves," *J. Acoust. Soc. Am. (JASA)*, vol. 45, no. 3, pp. 592–603, 1969.