Introduction to Ultrasonic Mid-Air Haptic Effects

Takayuki Hoshi

Abstract In this chapter, we discuss the basic physical principles of ultrasonic mid-air haptics. Our aim is to provide a holistic introduction to the technology, facilitate a better understanding of these principles, and help newcomers to join this exciting field of research in following the rest of this book. To that end, we have assumed a simplified and idealized situation and divide our discussion into four sub-topics: acoustic radiation pressure, phased array focusing, vibrotactile stimulation, and by-product audible sounds.

1   Introduction

Ultrasonic mid-air haptics is realized by exploiting several physical phenomena and control methods. It is necessary for researchers and engineers of this technology to understand those underlying principles, and such knowledge is also advisable for haptic designers, even if they are not from an engineering background. This chapter briefly introduces the principles of ultrasound mid-air haptics and acts as a gateway to the more advanced topics covered elsewhere in this book.

We will focus on four principles, leaving deeper discussions on specific aspects and application of this technology to other chapters or research papers. To grasp the essence, we assume the simplest conditions and describe the phenomena from a theoretical standpoint. The differences between theory and practice, e.g., due to approximations and simplification of the acoustic theory or due to the complexities of the hardware electronics, are not discussed in this chapter. The four topics we believe form a good starting point are as follows:

1. Acoustic radiation pressure: The force acting on the skin originates from the acoustic radiation pressure, a nonlinear effect caused by high-intensity sound waves. It is known that the acoustic radiation pressure is proportional to the acoustic energy density in front of the skin surface. We derive the acoustic

T. Hoshi (✉)
Pixie Dust Technologies, Inc., Tokyo, Japan
e-mail: star@pixiedusttech.com

radiation pressure by starting from the kinetic theory of gases to determine the origin.

2. Phased array focusing: Mid-air haptics is usually delivered by an array of hundreds of ultrasonic transducers. Although each transducer cannot radiate an intense ultrasonic wave, the phases of the transducers can be appropriately controlled to generate focal points as a result of the principle of superposition. We discuss the case of a single ultrasonic focal point and then briefly consider that of multiple focal points.

3. Vibrotactile stimulation: Human tactile perception is more sensitive to vibrations than to static pressure due to the underlying mechanoreceptors in the skin. In ultrasonic mid-air haptics, the focal point is usually amplitude-modulated (AM) to provide a vibrotactile sensation. Different modulation techniques exist and can be used to design different tactile feelings and experiences.

4. Audible sound radiation: The amplitude modulation of ultrasound waves can also cause unwanted audible sounds to be created. This is another nonlinear effect of high-intensity sound waves. The rapid movement of a focal point is another origin of audible sound, as the discontinuity of phase changes leads to fluctuations in the amplitude of the ultrasound. We consider methods for suppressing these noises.

2 Acoustic Radiation Pressure

2.1 Mathematical Expression

When an intense sound wave is blocked by an object, a force acts in the wave direction pushing the surface of the object. This effect is due to the acoustic radiation pressure, and it is one of the nonlinear effects of sound waves (Awatani 1955; Hasegawa et al. 2000). Iwamoto et al. first introduced this effect into mid-air haptics (Iwamoto et al. 2008) with the following explanation:

"The acoustic radiation pressure P [Pa] is described as

$$P = \alpha E = \alpha \frac{I}{c} = \alpha \frac{p^2}{\rho c^2}$$

where E [J/m^3] is the energy density of the ultrasound, I [W/m^2] is the sound power, c [m/s] is the sound speed, p [Pa] is the sound pressure of the ultrasound, and ρ [kg/m^3] is the density of the medium. α is a constant ranging from 1 to 2 depending on the reflection properties of the surface of the object. In case the surface of the object perfectly reflects the incident ultrasound, the value of α is 2, while if it absorbs the entire incident ultrasound, the value of α is 1. ... When the airborne ultrasound is applied on the surface of the skin, almost all the incident ultrasound is reflected."

(Note: p is the root mean square (RMS) value, that is, $1/\sqrt{2}$ of the peak amplitude [Pa] of a sinusoidal wave.)

The above explanation indicates that the acoustic radiation pressure is proportional to the square of the sound pressure. This means that the acoustic radiation pressure is too small to be felt on our skin when the sound pressure is small and becomes prominent when the sound pressure reaches a certain sound pressure level (SPL), such as approximately 140 dB or more. The perceptual thresholds for ultrasound are presented in Chap. "Ultrasound Exposure in Mid-Air Haptics". The sound pressure level (SPL) is a relative value of the RMS sound pressure p [Pa] to the RMS reference sound pressure $p_0 = 20\,\mu$Pa, and it is calculated through $20\log_{10}(p/p_0)$ [dB].

2.2 Derivation Based on Kinetic Theory of Gases

Our understanding of the phenomenon is as described above; however, it is difficult to grasp the physical intuition of it just by looking at the equation. We therefore take a step back and look at the kinetic theory of gases, and from there derive the acoustic radiation pressure.

We assume an ideal gas, which is a large number of identical submicroscopic particles, all of which are in constant, rapid, random motion. We consider the situation shown in Fig. 1 to derive the pressure on a wall whose area is S [m^2]. Although each particle has a different speed, we assume the average speed normal to the wall of the particle, v [m/s]. The total mass of the particles hitting and rebounding from the wall during time Δt [s] is calculated as a product of the density ρ and volume $Sv\Delta t$. Then, the impulse $F\Delta t$ [kg m/s] that the wall receives is calculated as a product of the total mass and the average speed, as follows:

$$F\Delta t = 2\rho Sv^2\Delta t$$

Here, "2" denotes a perfectly elastic collision, where v turns into $-v$ (i.e., the change in the average speed is $2v$) owing to the collision. From this equation, the atmospheric pressure P_0 [Pa] acting on the wall is obtained as follows:

$$P_0 = \frac{F}{S} = 2\rho v^2$$

Fig. 1 Illustration of a single molecule
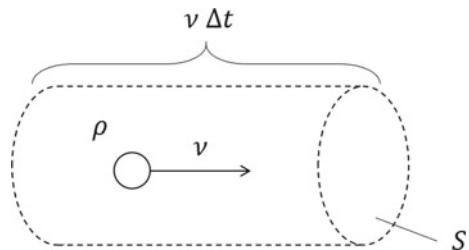
Next, we add a sound wave whose particle velocity is u [m/s], propagating normal to the wall. By replacing v in the above equation with $v + u$, the total pressure $P_{\text{tot}} = P_0 + P$ is given as follows:

$$P_{\text{tot}} = 2\rho(v + u)^2 = 2\rho v^2 + 4\rho vu + 2\rho u^2$$

Then, by taking the time average of P_{tot}, we can obtain an equation as follows:

$$\langle P_{\text{tot}} \rangle = \langle 2\rho v^2 \rangle + \langle 2\rho u^2 \rangle = P_0 + 2\rho u_{\text{rms}}^2$$

In the above, $\langle \cdot \rangle$ denotes the time average. The time average of $4\rho vu$ is zero, because $4\rho v$ is constant, and u is sinusoidal. The squared RMS value u_{rms}^2 gives the time-averaged squared value $\langle u^2 \rangle$ by definition. Finally, assuming a plane wave whose acoustic impedance is ρc, that is, $u = p/\rho c$, we have reached the same equation as explained in Iwamoto et al. (2008):

$$P = 2\frac{p_{\text{rms}}^2}{\rho c^2}$$

This final form is equal to $2E$, i.e., twice the acoustic energy density. This derivation process indicates that the origin of the acoustic radiation pressure is not the acoustic energy density itself, but rather the product of the momentum (containing u) and the number of particles hitting and rebounding from the wall (containing another u).

2.3 Diagonal Incidence

In the above derivation, we considered the case in which ultrasonic waves are incident normally on a flat surface. Here, we consider the case where they are incident diagonally, as discussed in Awatani (1955). In that study, it was shown that the radiation pressure normal to the wall was $(\cos\theta)^2$ times when the incident angle was θ. This was obtained by substituting $u\cos\theta$ with u in the above derivation.

Notably, if the width of the incident ultrasonic beam is limited to S, the total force applied by the diagonal incidence is multiplied by a factor of $\cos\theta$. This is because the surface area that receives the ultrasonic beam becomes larger (i.e., $S/\cos\theta$), as shown in Fig. 2. The radiation force is the product of the radiation pressure (i.e., $P(\cos\theta)^2$) and surface area. Therefore, we have a force of $PS\cos\theta$.

Fig. 2 Incident and reflected beams of ultrasound at the angles of incidence and reflection, θ. The surface area within the beam is $1/\cos\theta$ times of the cross-sectional area of the beam, S
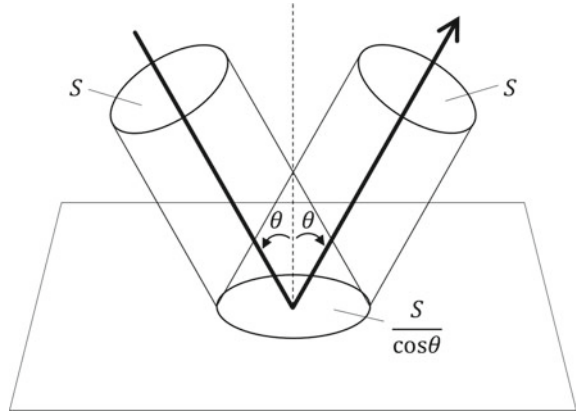
2.4 Acoustic Streaming

In ultrasonic mid-air haptics, not only the radiation pressure is felt but also a feeling of air flow. This effect is referred to as acoustic streaming, a nonlinear phenomenon of high-intensity sound waves. It is often assumed that the sound wave does not move the medium, and only the sound energy propagates through it when the amplitude is small. However, when the wave amplitude increases, a part of the acoustic energy acts as a driving force, and a flow in the medium is generated. This is initially a laminar flow, and then, it becomes a turbulent flow as it develops. Users of ultrasonic mid-air haptics may feel not only the sense of pressure owing to the radiation pressure, but also a sense of wind owing to acoustic streaming. Further details of this phenomenon are provided in Chap. "The Physical Principles of Arrays for Mid-Air Haptic Applications".

3 Phased Array Focusing

In the previous section, we demonstrated that high-intensity ultrasound can push against objects. The next issue we discuss concerns how to generate high-intensity ultrasound. Conventionally, a bolt-clamped Langevin-type transducer (BLT) has been used to generate high-intensity ultrasonic waves. For example, Ito achieved an SPL of 178 dB at the center of a focal point by combining a BLT with a properly designed concave reflector for converging the ultrasonic wave (Ito 2015). Another method of achieving a high-intensity ultrasonic wave in the air is to use a large number of transducers that cannot individually output such high-intensity ultrasonic waves. Together, these transducers can generate a high-intensity focal point by appropriately controlling their phases (Hoshi et al. 2010). For example, Hoshi et al. achieved an SPL of 162 dB at the center of a focal point with 285 ultrasonic transducers.

This was sufficiently intense to make people feel tactile sensations (Hoshi 2014). Furthermore, we can move the position of the focal point electronically by controlling the transducer phases. Therefore, the device that generates these focal points and control their respective positions is called a phased array.

Below, we describe the sound pressure distribution generated by a phased array in the case where a single focal point is generated. The first step toward multiple focal points is also introduced.

3.1 Focal Point

The spatial distribution of an ultrasonic focal point of pressure p [Pa] on the focal plane coordinates (x_f, y_f) at the focal length r [m] generated by $N \times N$ transducers arranged in a square lattice at interval d [m] is given by the following equation.

$$p(x_f, y_f) \approx \sqrt{2} p_r N^2 \frac{\text{sinc}\left(\frac{Ndv_x}{2}, \frac{Ndv_y}{2}\right)}{\text{sinc}\left(\frac{dv_x}{2}, \frac{dv_y}{2}\right)} e^{j\{\varphi(x_f, y_f) - \omega t\}}$$

Here, we have ignored the directivity of the transducer and assumed that a spherical wave is radiated from each transducer. Transducer directivity is considered in Chaps. "The Physical Principles of Arrays for Mid-Air Haptic Applications" and "Prototyping Airborne Ultrasonic Arrays". The parameter definitions and some derivations are provided in Appendix A. The following are notable points from the above equation of $p(x_f, y_f)$.

- The sound pressure at the center of the focal point is the product of the sound pressure traveled from a single transducer p_r and the number of transducers N^2.
- The spatial distribution of the ultrasound on the focal plane follows the sinc function, i.e., $\text{sinc}(x, y) \equiv \sin(x) \sin(y)/xy$.

The focal point as represented by the sinc function described above has a spatial distribution of the ultrasound, as shown in Fig. 3. The diameter w [m] of the region with the largest amplitude (main lobe) is expressed by the following equation using three parameters: the wavelength λ [m] of the ultrasonic wave, size of the phased array Nd, and focal length r:

$$w = 2\lambda \frac{r}{Nd}$$

The above equation is derived from the condition that the sinc function of the numerator of the equation of the sound pressure distribution $p(x_f, y_f)$ first becomes zero. From this equation, it can be seen that if the phased array is too small or the focal length is too large, the focal point will be blurred, i.e., the ultrasonic waves will be dispersed, and a high-intensity ultrasound will not be obtained. This is determined

Fig. 3 Spatial distribution of ultrasound (absolute value) on the focal plane generated by a square-shaped phased array (close-up shot: $40 \times 40\,\text{mm}^2$). The array size is 170 mm, and the focal length is 200 mm
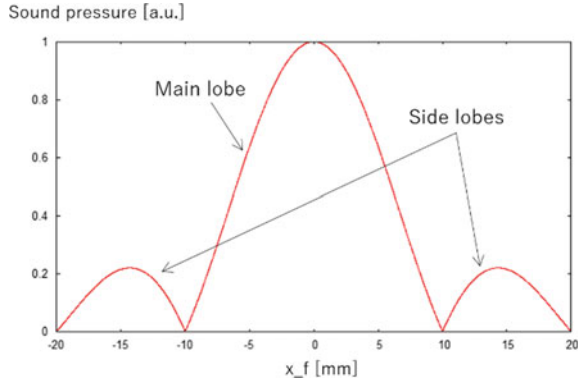
(a) 1D cross-sectional plot of focal point ($y_\text{f} = 0$).
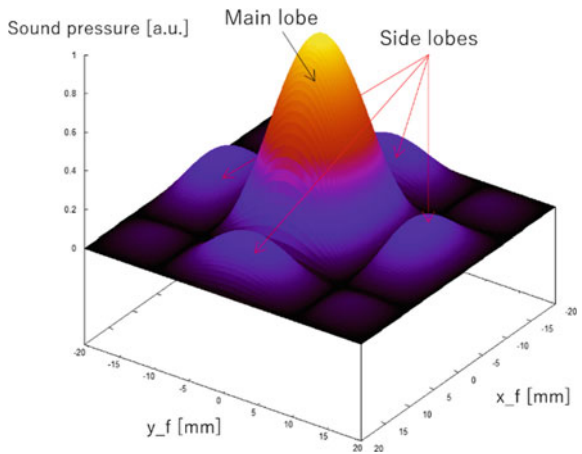
(b) 2D plot of focal point.

by the interference of the waves and is related to numerical values in the field of optics such as the diffraction limit, F value, and the numerical aperture.

Although the main lobe is surrounded by multiple side lobes (see Fig. 3), it is called a focal point because the side lobes are often too weak to be felt and are therefore ignored. For example, the diameter of the focal point (the full width at the zero intensity of the main lobe) is calculated as 20 mm when the wavelength is 8.5 mm (40 kHz ultrasonic wave in air), the length of one side of the square-shaped phased array is 170 mm, and the focal length is 200 mm. Notably, the area in which the radiation pressure is sufficiently strong to evoke a tactile sensation may be narrower than the calculated diameter.

The shape of the focal point here is the sinc function, because the shape of the phased array is square. These shapes are in the relationship of a Fourier transform, and the shape of the focal point changes if the shape of the phased array changes. For example, we have a focal point represented by the Airy pattern when we use a
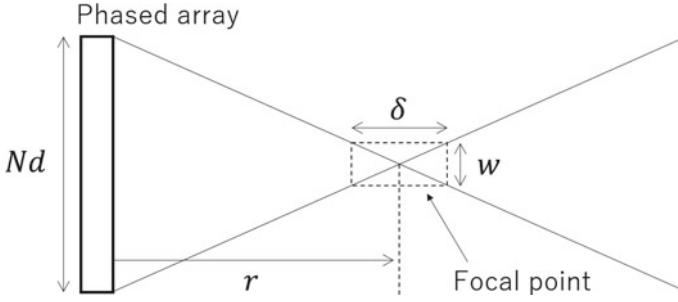
Fig. 4 Width and depth of the focal point © IEEE. Reprinted, with permission, from Hoshi et al. (2010)

circular phased array. As explained in the field of lens optics, the radius of the Airy pattern is given by $1.22\lambda r/D$, where D is the diameter of the circular wave source. Thus, the diameters of the main lobes produced by the square and circular phased arrays are $2\lambda r/Nd$ and $2.44\lambda r/D$, respectively. These are not the same, but are not significantly different if the diameters of the arrays are similar, that is, if $Nd = D$. As another example, in the case of a hexagonal phased array, the shape of the focal point is radial in six directions, i.e., it will exhibit 6 smaller side lobes.

As shown above, a focal point is not truly a point, and it has both a width and depth. This means that the users can feel tactile sensations on their hand within the focal depth, although the sound pressure gradually decreases as the distance from the peak increases. This alleviates the accuracy required for the depth control of the focal position. The peak exists approximately a few centimeters before the focal point (target) in the Z direction, owing to the limited size of the phased array.

Here, we derive the focal depth δ [m], assuming the situation shown in Fig. 4. The relationship among the array size, focal length, width, and half depth of the focal point is derived based on the similarity relationship between triangles, as follows:

$$\frac{Nd}{r} = \frac{w}{\delta/2}$$

By substituting $w = 2\lambda r/Nd$ as obtained in Sect. 3.1, the focal depth δ is given as follows:

$$\delta = \frac{4\lambda r^2}{N^2 d^2}$$

For example, $\delta = 47$ mm when the wavelength of the ultrasonic wave is 8.5 mm, the length of one side of the square-shaped phased array is 170 mm, and the focal length is 200 mm.
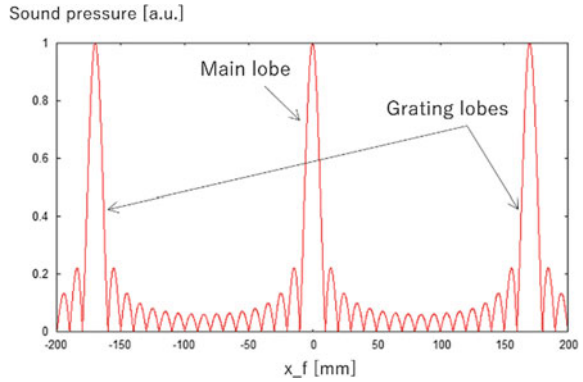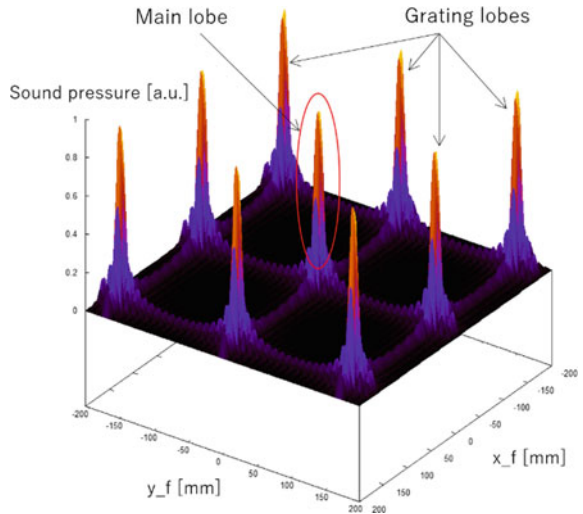
3.2 Grating Lobes

Attention should be paid to when the interval between the centers of neighboring transducers is larger than the wavelength of the ultrasonic waves. This is because regions called grating lobes are generated having a large amplitude comparable to the focal point (main lobe) (see Fig. 5). The distance l [m] between the main lobe and the first grating lobe is given by the following equation,

$$l = \lambda \frac{r}{d}$$

Fig. 5 Spatial distribution of ultrasound (absolute value) on the focal plane generated by a square-shaped phased array (long shot: $400 \times 400\,\text{mm}^2$). The transducer interval is 10 mm, and the focal length is 200 mm.

(a) 1D cross-sectional plot of main and grating lobes $(y_\text{f} = 0)$.

(b) 2D plot of main and grating lobes.

This is derived from the condition that the sinc function of the denominator of the equation of the sound pressure distribution $p(x_f, y_f)$ first becomes zero. For example, the grating lobes are generated at $l = 170$ mm in the X and Y directions, that is, in the direction of 40° with respect to the main lobe when the transducers are arranged at an interval $d = 10$ mm, the wavelength of the ultrasonic wave is 8.5 mm, and the focal length is 200 mm. This angle (40°) is determined by the ratio of the transducer interval to the ultrasonic wavelength and becomes larger as the spacing decreases. They finally disappear when the interval becomes shorter than the wavelength, because the condition for the generation of the grating lobes is no longer fulfilled. This disappearance of the grating lobes cannot be explained by l as shown above, because l was derived under the paraxial and Fresnel approximations.

In practice, the grating robes are weaker than the main lobe, owing to the directivity of the ultrasonic transducers that we have ignored in our discussion. An actual transducer has a half-angle at half maximum, such as 50°. That is, the sound pressure radiated in the direction of 50° is 6 dB smaller than that in the direction of 0°. From this, it is predicted that the sound pressure at the grating lobe appearing in the direction of 40° is less than the sound pressure of the focal point by nearly 6 dB, and even less in the direction of more than 40°. Furthermore, the radiation force on the hand generated by the grating lobe at the focal plane also suffers from a reduction owing to the incident angle as discussed in Sect. 2.3, i.e., approximately −2 dB for 40°. A further reduction in the grating lobe importance is due to the squaring of p in order to calculate the acoustic radiation pressure P.

The grating lobes shown above originated from the periodic placement of the transducers and the inter-transducer distances. It is known that arrays whose transducers are not arranged in a periodic manner mitigate the grating lobes (for example, see Price and Long (2018)). However, such an arrangement reduces density of transducers that can be laced on a PCB leading to a reduction in the output ultrasonic power. Thus, a trade-off exists between high pressure and grating lobes.

3.3 Multiple Focal Points

Although an example of generating a single focal point is shown in the above calculation, there are also algorithms for generating multiple focal points. In the most commonly used case, a matrix equation is solved to determine adequate phases of the transducers (Carter et al. 2013; Long et al. 2014). This is an important calculation not only for mid-air haptics, but also for acoustic levitation (for example, see (Morales et al. 2019)). The specific calculation methods are discussed further in Chap. "Sound-Field Creation for Haptic Reproduction", and only the essence of the formulation is introduced here.

We consider the situation shown in Fig. 6. Here, N transducers are arranged on the phased array, and M focal points are to be generated. The controllable parameters of the n-th transducer are the amplitude x_n and phase α_n of the driving signal. Then, all of the initial sound waves radiated from the transducers can be represented as

Fig. 6 Formulation of multiple focal points. Here, as an example, N transducers generate two focal points

Focal points

Phased array
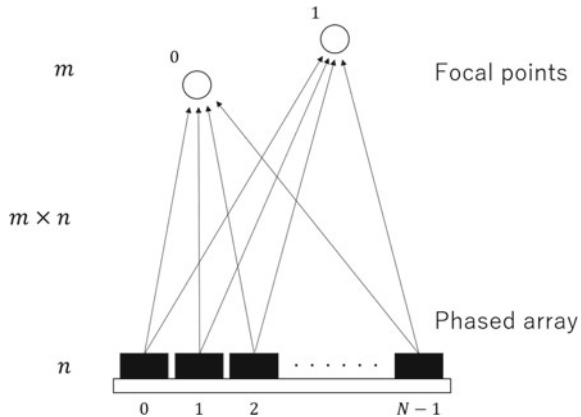
a vector $x = \{x_n \mathrm{e}^{\mathrm{j}\alpha_n}\}$. Similarly, the parameters of the m-th focal point are the amplitude b_m and phase β_m of the sound pressure giving the vector $b = \{b_m \mathrm{e}^{\mathrm{j}\beta_m}\}$. The propagation of a sound wave is represented as $\mathrm{e}^{\mathrm{j}kr_{m,n}}/r_{m,n}$, where $r_{m,n}$ [m] is the distance from the n-th transducer to the m-th focal point. Assuming a propagation matrix $A = \{\mathrm{e}^{\mathrm{j}kr_{m,n}}/r_{m,n}\}$, we have a linear equation that maps x onto b via:

$$b = Ax$$

Now the problem comes down to solving this matrix equation and how to obtain the inverse matrix of A efficiently and under some constraints and conditions. One simple solution is using the pseudo-inverse matrix $A^+ = \left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}$, where superscripts T and -1 denote the transpose and the inverse, respectively. Then, x is obtained as $x = A^+b$. However, this solution has not been optimized in terms of the overall spatial distribution of ultrasound, i.e., only the amplitudes at the focal points are designed, and the other areas are not considered. Therefore, additional processing, such as regularization and iterative calculations, may be required to brush up the solution.

4 Vibrotactile Stimulation

In the case of most ultrasonic mid-air haptic devices, the force applied on a reflective surface by the high-intensity ultrasound is slightly greater than 10 mN within an area of several centimeters in diameter. If this is continuously applied to the skin surface, the nerves of the mechanoreceptors under the skin will adapt, and the tactile sensation will no longer be felt. AM of the ultrasound is usually used to provide a vibrotactile stimulation (Fig. 7) which can be more efficiently sensed by the receptors embedded in our skin. Recently, different modulation methods other than AM have also been
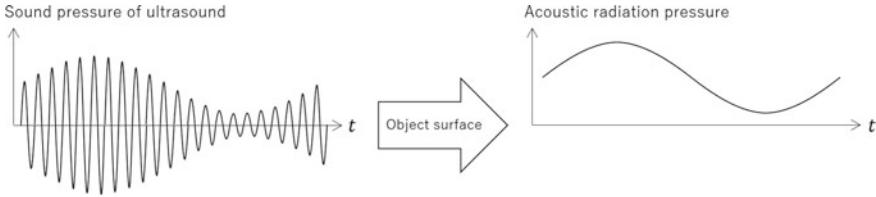
Fig. 7 Acoustic radiation pressure acting on the object surface produced by the amplitude modulation (AM) ultrasound

proposed; these are discussed further in Chap. "Modulation Methods for Ultrasound Midair Haptics".

The modulation frequency and waveform used at a focal point can affect the texture of the tactile sensations related to human tactile perception characteristics. Thus, the frequency content of any vibrotactile stimulus needs to match the characteristics of the human tactile sensory channels, which are approximately in the range of 10–500 Hz (Bolanowski et al. 1988). Therefore, a modulation of 100–200 Hz is usually used for mid-air haptics. Notably, the spatial resolution and the perceived tactile sensation differs depending on the stimulus frequency (Vallbo and Johansson 1984). For example, the focal point at 200 Hz feels larger and more blurred than one at 50 Hz (Hoshi 2015). There is many degrees-of-freedom and many trade-offs that one needs to consider when designing modulation techniques as they will each induce a different tactile sensation. This is an active research space where different authors have begun to build up a library of mid-air haptic sensations.

5   Audible Sound Radiation

Audible sounds originating from AM ultrasound have been reported. Even though the ultrasound itself is inaudible, such audible sounds can sometimes worsen the experience. This side effect is caused because the tactile frequency range (approximately 10–500 Hz) of the mid-air haptic stimulus overlaps with the audible frequency range (approximately 20 Hz to 20 kHz). Here, the mechanism of the audible sound generation is explained, and methods for suppressing noise are introduced.

5.1   Self-demodulation

The audible frequency range varies from person to person, but it is generally from 20 Hz to 20 kHz. Sounds higher than 20 kHz are usually called ultrasound. For example, 40 kHz is twice as high as the audible range; hence, the human ear usually
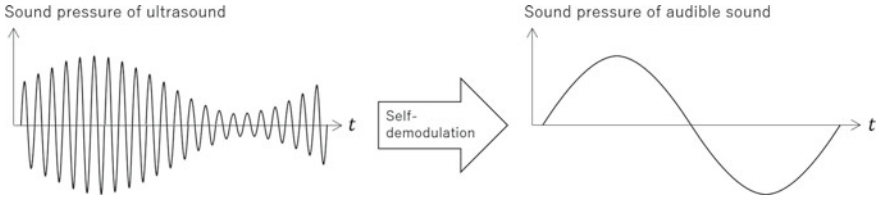
Fig. 8 Audible sound radiated from the AM ultrasound

hears nothing from a constant 40 kHz signal. However, it is known that when ultrasound is more intense than a certain sound pressure level and the ultrasound fluctuates at frequencies in the audible range, the fluctuation is radiated as an audible sound. In other words, as the modulated ultrasonic waves propagate through the air, they are demodulated by the nonlinearity of the air as seen in Fig. 8. The result is that the space in which the ultrasonic waves exist behaves as a source of audible sound; a kind of self-demodulating effect.

The audible sound p_s [Pa] produced by self-demodulation can be expressed by a differential equation, as follows (Yoneyama et al. 1983):

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)p_s = -\frac{\beta}{\rho c^2}\frac{\partial^2}{\partial t^2}p^2$$

Here, ρ is the density of air, c is the speed of sound, and β is the nonlinear parameter used in nonlinear acoustics ($\beta = 1.2$ for air). The left-hand side of this differential equation is the wave equation for audible sound. The right side indicates that the time change of the ultrasound acts as a driving force. It is a nonlinear phenomenon that is driven by the time-second derivative of the square of the sound pressure of the ultrasound p signal.

Here, we will briefly examine this phenomenon based on mathematical formulas. In the case of AM by a sinusoidal wave envelope, the modulated ultrasound is expressed as follows:

$$p = (P_c + P_m \cos\omega_m t)\cos\omega_c t$$
$$= P_c \cos\omega_c t + \frac{P_m}{2}\cos(\omega_c + \omega_m)t + \frac{P_m}{2}\cos(\omega_c - \omega_m)t$$

In the above, P and ω are the amplitude and angular frequency, respectively, and the subscripts c and m denote the "carrier" (i.e., usually 40 kHz) and "modulation" (i.e., the frequency lower than the carrier wave). As the second line shows, the resulting AM signal p has three frequency components, ω_c, $\omega_c + \omega_m$, and $\omega_c - \omega_m$, shown in Fig. 9, left. We can then investigate how this signal acts as a sound source by substituting the modulated ultrasound p into the right side of the self-demodulation differential equation from above:
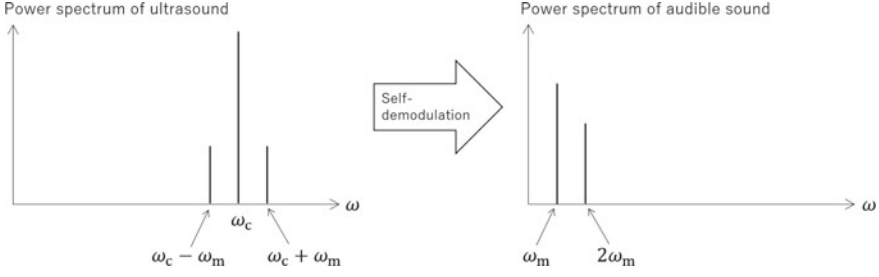
Fig. 9 Power spectrums of the AM ultrasound and radiated audible sound

$$-\frac{\beta}{\rho c^2}\frac{\partial^2}{\partial t^2}p^2 = -\frac{\beta}{\rho c^2}\frac{\partial^2}{\partial t^2}(P_c + P_m \cos\omega_m t)\cos\omega_c t$$

$$= \frac{\beta}{\rho c^2}\Big\{ P_c P_m \omega_m^2 \cos\omega_m t + P_m^2 \omega_m^2 \cos 2\omega_m t$$

$$+ \left(2P_c^2\omega_c^2 + P_m^2\omega_m^2\right)\cos 2\omega_c t$$

$$+ \frac{4P_c P_m \omega_c^2 + P_c P_m \omega_m^2}{2}[\cos(2\omega_c + \omega_m)t + \cos(2\omega_c - \omega_m)t]$$

$$+ 2P_c P_m \omega_c \omega_m[\cos(2\omega_c + \omega_m)t - \cos(2\omega_c - \omega_m)t]$$

$$+ \frac{P_m^2\omega_c^2 + P_m^2\omega_m^2}{2}[\cos(2\omega_c + 2\omega_m)t + \cos(2\omega_c - 2\omega_m)t]$$

$$+ \frac{P_m^2\omega_c\omega_m}{2}[\cos(2\omega_c + 2\omega_m)t - \cos(2\omega_c - 2\omega_m)t]\Big\}$$

As shown above, various frequency components are included in the driving force. When we omit the frequency components in the ultrasonic range, we have the frequency components of ω_m and $2\omega_m$ in the audible range (Fig. 9, right). It is also shown that the radiated sound has a frequency dependency of ω_m^2, that is, the lower frequency is less radiated (-12 dB when the modulation frequency is half). Notably, a higher frequency is also less radiated, owing to the narrow resonant frequency band of the ultrasonic transducer (e.g., -6 dB at approximately 1.5 kHz away from the resonant frequency).

Possible ways to reduce this noise include making the modulation waveform as close to a sinusoidal wave as possible to reduce the extra frequency components, and setting the modulation frequency lower than the audible range, so that users cannot hear the by-product sound. Recently, it has been reported that the lateral modulation (spatial movement of the focal point on the skin surface) is less noisy than AM (Suzuki et al. 2020). This is an idea based on using spatial control instead of temporal control and on employing a gradual phase change to reduce the noise, as discussed in the next subsection.

5.2 *Movement of Focal Point*

The movement of the focal point is achieved by appropriately changing the phase command value of each transducer. However, the transducer being a resonant system, it exhibits a transient response when its input signal is stepwise switched. Specifically, it can be seen from simulating an equivalent circuit of an ultrasonic transducer that the pressure amplitude experiences a significant dip because of this phase switching during a movement of the focal point. This change in amplitude produces an audible sound. While the audible sound generated by one transducer is small, it becomes significant and perceivable when hundreds of transducers are synchronized in the phased array. To reduce this effect, it is sufficient that the phase change rate is small. Thus, a method of updating the position change of the focal position with a high spatial resolution and high update rate should be considered. Although this is effective when the intended focal movement trajectory is nearly continuous, the effect of this method is limited when the focal point hops between discrete positions. To suppress any noise artifact in that case, a method of gradually changing the phase has been proposed (Hoshi 2016; Hoshi, 2020), and a noise suppression of approximately 10 dB was achieved in that case.

Here, we observe the effect of gradually changing the phase of the driving signal by simulation. A resonant electric circuit is modeled, and three types of driving signals of the resonance frequency (here, 40 kHz) are input into the circuit. Although the phase change is the same ($\pi/4$) for all three situations, the transitions are different: the phase changes instantly, gradually over seven cycles, and gradually over 14 cycles. The results are shown in Fig. 10, where the amplitude change occurs shortly after
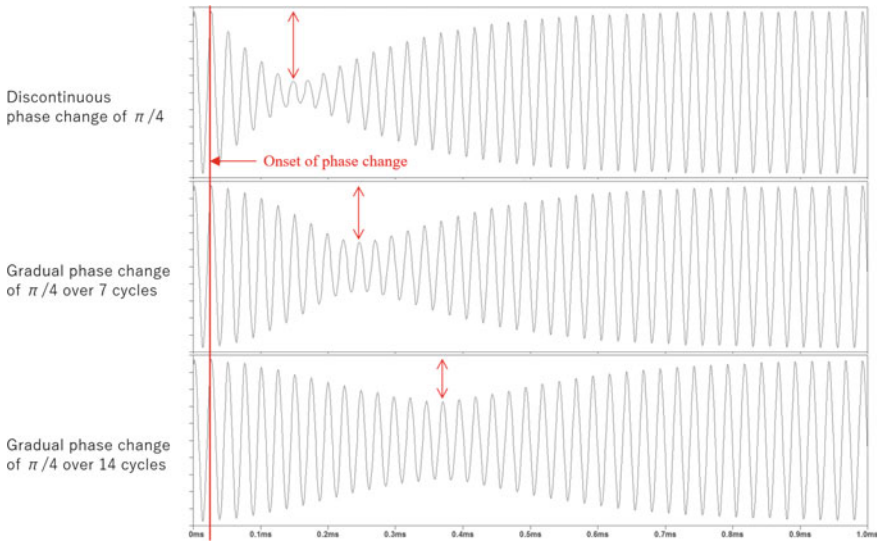

Fig. 10 Simulation of amplitude changes induced by discontinuous or gradual phase changes

the phase change, and the amplitude change decreases as the phase changes slowly. This control method takes a longer time than an instant phase change; however, the transient period is approximately 1.0 ms at the maximum, and it is therefore negligible for the purposes of mid-air haptics.

6   Conclusion

Multiple physical phenomena are involved in the generation process of ultrasonic mid-air haptic sensations, such as the acoustic radiation pressure that impinges the skin surface, phased array focusing algorithms based on interference, modulation techniques to induce a vibrotactile effect, and self-demodulation that radiates audible sound. These phenomena are usually discussed in the form of mathematical equations. While such equations are useful for diving into deeper discussions and engineering, they are not easy to understand and pose an entry barrier to newcomers of the field of mid-air haptics. To that end, we have attempted to explain the principles of ultrasonic mid-air haptics using simplifying assumptions, visual representations, and basic physical argumentation.

First, the origin of the acoustic radiation pressure was shown, based on the kinetic theory of gases, assuming an ideal gas and plane ultrasonic wave. The effect of the incidence angle was discussed. Another effect known as acoustic streaming was also introduced. Please refer to Chap. "The Physical Principles of Arrays for Mid-Air Haptic Applications" for further discussion.

Second, phased array focusing was explained using a square-shaped array, based on the paraxial approximation and the Fresnel approximation, while neglecting the directivity of the ultrasonic transducer. Both the main lobe and grating lobes were shown. A conceptual formulation was also provided for calculating adequate amplitudes and phases for generating multiple focal points. Please see Chaps. "The Physical Principles of Arrays for Mid-Air Haptic Applications" and "Prototyping Airborne Ultrasonic Arrays" for the generation of the acoustic field, and Chap. "Multiunit Phased Array System for Flexible Workspace" for advances in phased arrays.

Third, it was noted that vibrotactile stimulation was used to effectively provide the tactile sensations based on the human tactile perception characteristics and that such stimulation is usually provided by AM. If interested in the variety of modulation methods, please refer to Chap. "Modulation Methods for Ultrasound Midair Haptics".

Fourth, the radiation of audible sounds was introduced. Two sources were discussed, the amplitude modulation frequency and the sudden phase change when moving the ultrasonic focal point. The suppression of these audible noises was also discussed.

Finally, we hope that the readers of this chapter, and indeed of this book, will use this text as an introductory reference for ultrasonic mid-air haptics. Moreover, we hope that this crash course introduction can further motivate multidisciplinary

research in hardware, software, waveform, and haptic experience design, thus accelerating the advancement of mid-air haptic technology and its uptake in novel human computer interaction applications and use cases.

Appendix

A. Derivation of Spatial Distribution of Ultrasound on Focal Plane

The process by which a phased array forms a single ultrasonic focal point can be expressed using the following mathematical formulas. We assume the coordinate system shown in Fig. 11. The XY coordinate is on the surface of the phased array (x, y). The focal plane, (x_f, y_f), is separated by the distance r [m] from the surface. Let the center coordinate of the m-th and n-th transducers in the X and Y directions, respectively, be $(x_m, y_n, 0)$, and the center coordinate of the focal point be (x_c, y_c, r).

Here, we ignore the directivity of the transducer, and assume that a spherical wave is radiated from each transducer with an appropriate phase to generate a single focal point. The RMS ultrasonic sound pressure p_r [Pa] (inversely proportional to r) is produced on the focal plane by each transducer and is a paraxial approximation. A square-shaped phased array is assumed; this is the shape of the majority of current ultrasonic mid-air haptic devices. The N transducers are lined up in both the X and Y directions. At this time, the distribution of ultrasonic sound pressure generated on the focal plane is given as follows:

Fig. 11 Coordinate system for the formulation of phased array focusing © IEEE. Reprinted, with permission, from Hoshi et al. (2010)
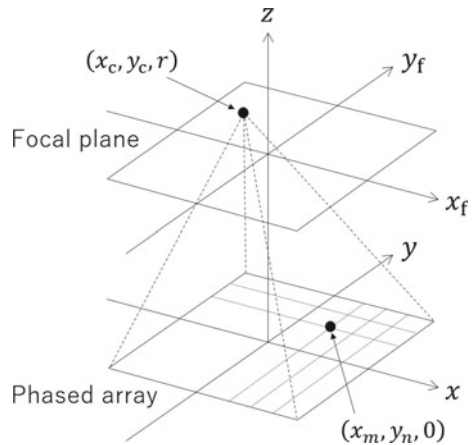
$$p(x_f, y_f) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \sqrt{2}\, p_r e^{-jkr'} e^{j(kr''-\omega t)}$$

$$\approx \sqrt{2}\, p_r e^{-j\omega t} e^{j\{\frac{k}{2r}(x_f^2+y_f^2-x_c^2-y_c^2)-\xi(v_x+v_y)\}} \sum_{m=0}^{N-1} e^{-jdmv_x} \sum_{n=0}^{N-1} e^{-jdnv_y}$$

$$= \sqrt{2}\, p_r e^{-j\omega t} e^{j\{\frac{k}{2r}(x_f^2+y_f^2-x_c^2-y_c^2)-\xi(v_x+v_y)\}} \left(\frac{1-e^{-jNdv_x}}{1-e^{-jdv_x}}\right)\left(\frac{1-e^{-jNdv_y}}{1-e^{-jdv_y}}\right)$$

$$= \sqrt{2}\, p_r e^{-j\omega t} e^{j\{\frac{k}{2r}(x_f^2+y_f^2-x_c^2-y_c^2)-\xi(v_x+v_y)\}} e^{-j\frac{(N-1)d}{2}(v_x+v_y)}$$

$$\times \left(\frac{e^{j\frac{Ndv_x}{2}}-e^{-j\frac{Ndv_x}{2}}}{e^{j\frac{dv_x}{2}}-e^{-j\frac{dv_x}{2}}}\right)\left(\frac{e^{j\frac{Ndv_y}{2}}-e^{-j\frac{Ndv_y}{2}}}{e^{j\frac{dv_y}{2}}-e^{-j\frac{dv_y}{2}}}\right)$$

$$= \sqrt{2}\, p_r\, e^{j\{\varphi(x_f, y_f)-\omega t\}} \frac{\sin\frac{Ndv_x}{2}}{\sin\frac{dv_x}{2}} \frac{\sin\frac{Ndv_y}{2}}{\sin\frac{dv_y}{2}}$$

$$= \sqrt{2}\, p_r\, e^{j\{\varphi(x_f, y_f)-\omega t\}} \left(\frac{\frac{Ndv_x}{2}\frac{Ndv_y}{2}}{\frac{dv_x}{2}\frac{dv_y}{2}}\right) \frac{\sin\left(\frac{Ndv_x}{2}\right)\sin\left(\frac{Ndv_y}{2}\right)/\left(\frac{Ndv_x}{2}\frac{Ndv_y}{2}\right)}{\sin\left(\frac{dv_x}{2}\right)\sin\left(\frac{dv_y}{2}\right)/\left(\frac{dv_x}{2}\frac{dv_y}{2}\right)}$$

$$= \sqrt{2}\, p_r\, N^2 \frac{\mathrm{sinc}\left(\frac{Ndv_x}{2}, \frac{Ndv_y}{2}\right)}{\mathrm{sinc}\left(\frac{dv_x}{2}, \frac{dv_y}{2}\right)} e^{j\{\varphi(x_f, y_f)-\omega t\}}$$

In the above, j is an imaginary unit. k [rad/m] and ω [rad/s] are the wavenumber and angular frequency, respectively. t [s] is the time. The definition of the sinc function is $\mathrm{sinc}(x, y) \equiv \sin(x)\sin(y)/xy$.

The first line represents the operation of multiplying the spherical wave coming from a transducer $\sqrt{2}p_r e^{j(kr''-\omega t)}$ by the phase control factor to focus the waves $e^{-jkr'}$ and then adding them together. In the second line, the Fresnel approximation is applied to r' and r'', where r' [m] is the distance from the m-th row and n-th column transducer to the focal point, and r'' [m] is the distance from the transducer to the arbitrary position (x_f, y_f) on the focal plane.

$$r' \equiv \sqrt{(x_m-x_c)^2 + (y_n-y_c)^2 + r^2}$$

$$\approx r + \frac{(x_m-x_c)^2 + (y_n-y_c)^2}{2r}$$

$$r'' \equiv \sqrt{(x_m-x_f)^2 + (y_n-y_f)^2 + r^2}$$

$$\approx r + \frac{(x_m-x_f)^2 + (y_n-y_f)^2}{2r}$$

Furthermore, changes of the variables are applied as follows:

$$v_x \equiv \frac{k}{r}(x_{\mathrm f} - x_{\mathrm c})$$

$$v_y \equiv \frac{k}{r}(y_{\mathrm f} - y_{\mathrm c})$$

Then, the positions of the m-th and n-th transducers are written in another form, that is, $(x_m, y_n) = (md + \xi, nd + \xi)$, using the interval between the centers of neighboring transducers d [m] and an offset $\xi = -(N-1)d/2$. The third line is obtained by the summation formula of the geometric series, i.e., $\sum_{n=0}^{N-1} \alpha^n = (1 - \alpha^N)/(1 - \alpha)$. The fourth line is the middle of the formula transformation to the fifth line. By extracting $\mathrm{e}^{-\mathrm{j}\alpha/2}$ from $(1 - \mathrm{e}^{-\mathrm{j}\alpha})$, a form of $(\mathrm{e}^{\mathrm{j}\alpha/2} - \mathrm{e}^{-\mathrm{j}\alpha/2})$ is obtained. The fifth line is derived because $(\mathrm{e}^{\mathrm{j}\alpha/2} - \mathrm{e}^{-\mathrm{j}\alpha/2}) = 2\mathrm{j}\sin(\alpha/2)$. $\varphi(x_{\mathrm f}, y_{\mathrm f})$ is the phase delay and depends on the position on the focal plane:

$$\varphi(x_{\mathrm f}, y_{\mathrm f}) \equiv \frac{k}{2r}(x_{\mathrm f}^2 + y_{\mathrm f}^2 - x_{\mathrm c}^2 - y_{\mathrm c}^2) - \left\{ \xi + \frac{(N-1)d}{2} \right\}(v_x + v_y)$$

The sixth line is again the middle of the formula transformation to the seventh line. By extracting $\alpha/2$ from $\sin(\alpha/2)$, a form of $\sin(\alpha/2)/(\alpha/2)$ is obtained; thus, the sinc function finally appears in the seventh line.

References

Awatani J (1955) Studies on acoustic radiation pressure. I. (General considerations). J Acoust Soc Am 27(2):278–281

Bolanowski SJ Jr, Gescheider GA, Verrillo RT, Checkosky CM (1988) Four channels mediate the mechanical aspects of touch. J Acoust Soc Am 84(5):1680–1694

Carter T, Seah SA, Long B, Drinkwater B, Subramanian S (2013) UltraHaptics: multi-point haptic feedback for touch surfaces. In: Proceedings of 26th annual ACM symposium on user interface software and technology (UIST'13), pp 505–514

Hasegawa T, Kido T, Iizuka T, Matsuoka C (2000) A general theory of Rayleigh and Langevin radiation pressures. Acoust Sci Technol 21:145–152

Hoshi T Low-noise ultrasonic wave focusing apparatus. U.S. Patent, US20170144190A1, 2020-02-25

Hoshi T, Takahashi M, Iwamoto T, Shinoda H (2010) Noncontact tactile display based on radiation pressure of airborne ultrasound. IEEE Trans Haptics 3(3):155–165

Hoshi T (2014) Noncontact tactile display using airborne ultrasound. In: Proceedings of 21st international display workshops, pp 1529–1533

Hoshi T (2015) Variation of tactile feelings of focused ultrasound: modulation frequency and hand movement. In: Proceedings of ACM SIGGRAPH Asia 2015, workshop on haptic media and contents design. Article 18

Hoshi T (2016) Gradual phase shift to suppress noise from airborne ultrasound tactile display. In: ACM CHI 2016 workshop on mid-air haptics and displays: systems for un-instrumented mid-air interactions, session 2: provide visual and haptic feedback

Ito Y (2015) High-intensity aerial ultrasonic source with a stripe-mode vibrating plate for improving convergence capability. Acoust Sci Technol 36:216–224

Iwamoto T, Tatezono M, Shinoda H (2008) Non-contact method for producing tactile sensation using airborne ultrasound. In: Haptics: perception, devices and scenarios: 6th international conference, Eurohaptics 2008 proceedings (lecture notes in computer science), pp 504–513

Long B, Seah SA, Carter T, Subramanian S (2014) Rendering volumetric haptic shapes in mid-air using ultrasound. ACM Trans Graph 33:181:1–181:10

Morales R, Marzo A, Subramanian S, Martínez D (2019) LeviProps: animating levitated optimized fabric structures using holographic acoustic tweezers. In: Proceedings of 32nd annual ACM symposium on user interface software and technology (UIST'19), pp 651–661

Price A, Long B (2018) Fibonacci spiral arranged ultrasound phased array for mid-air haptics. In: Proceedings of IEEE international ultrasonics symposium (IUS), pp 1–4

Suzuki S, Fujiwara M, Makino Y, Shinoda H (2020) Reducing amplitude fluctuation by gradual phase shift in mid-air ultrasound haptics. IEEE Trans Haptics 13(1):87–93

Vallbo ÅB, Johansson RS (1984) Properties of cutaneous mechanoreceptors in the human hand related to touch sensation. Hum Neurobiol 3:3–14

Yoneyama M, Fujimoto J, Kawamo Y, Sasabe S (1983) The audio spotlight: an application of nonlinear interaction of sound waves to a new type of loudspeaker design. J Acoust Soc Am 73(5):1532–1536