

Шаг 1 - 3DoF или 6DoF? (это решает всё)

3DoF	Только вращение головы: yaw, pitch, roll. Система вращает фиксированное звуковое поле. Дёшево. 360-видео, сидячий VR, фиксированный AR.
6DoF	Добавляет положение тела: x, y, z. Нужно пересчитывать дистанцию, направление, перекрытие при ходьбе. Это физическая симуляция. Нужен для VR в полный рост.
Правило	Степени свободы звука = степени свободы видео. 6DoF-видео + 3DoF-звук = звук не из того места = сцена рушится.

Шаг 2 - выбор SDK по платформе

Браузер / кроссплатформа: Web Audio API (PannerNode, panningModel HRTF) или Google Resonance Audio.

Meta Quest: Meta XR Audio SDK (HRTF + амбисоника + акустика помещения; заморожен на v85.0, 2025).

Apple Vision Pro / iOS: Apple PHASE (прямой путь, ранние отражения, поздняя реверберация; ReverbComponent в visionOS).

Физически точная VR-игра: Valve Steam Audio (моделирование отражений, TOA, опционально GPU).

Голосовой звонок с пространственным звуком: 3GPP IVAS (Release 18), когда устройства начнут поддерживать.

Два стандарта, которые только что вышли

MPEG-I Immersive Audio (ISO/IEC 23090-4:2025, ноябрь 2025): 6DoF-стриминг. Реверберация, отражения, перекрытие, дифракция, эффект Доплера.

3GPP IVAS (Release 18): пространственный звук в звонках, 13,2-512 кбит/с; 3-4 объекта при 24,4 / 32 кбит/с.

Шаг 3 - расчёт битрейта (каналов = (N+1) в квадрате)

FOA (порядок 1) = 4 канала. При 64 кбит/с на канал = ~256 кбит/с. Минимум для 'звук с любой стороны'.

TOA (порядок 3) = 16 каналов. При 64 кбит/с на канал = ~1 Мбит/с. В 4 раза больше бит ради точности.

Стримьте FOA, если контенту не нужна точность высокого порядка. (Реальные кодеры считают совместно - суммы ниже.)

Чек-лист перед запуском

Степени свободы звука совпадают с видео (оба 3DoF или оба 6DoF).

- ☐ Акустика помещения ВКЛ для любого опыта с передвижением - ранние отражения + реверберация, не только HRTF.

Слежение за головой проверено целиком на реальном устройстве; задержка движение-звук низкая.

- ☐ SDK выбран по целевой платформе, а не по привычке.
- ☐ Порядок амбисоники подобран под продукт; head-locked стерео для невращающегося нарратива.
- ☐