Notes on Acoustic Black Holes

Bryan W. Roberts

16 March 2021, CBG Part iii

1. Introduction

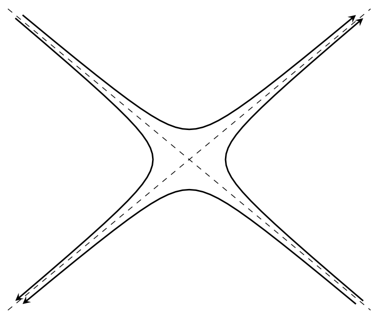

Figure 1.

It is commonly suggested that there are at least three interpretations of the diagram in Figure 1 above, which can be summarised as follows.

(1) **Flat spacetime:** Minkowski spacetime points in Euclidean coordinates;
(2) **Black hole:** A Schwarzschild black hole in Kruskal coordinates;
(3) **Supersonic fluid:** A continuous accelerating fluid; the dotted lines represent the sound-horizon at which the fluid moves at the speed of sound, and the solid lines are symmetries of the fluid's equation of motion.

The latter leads to the hope that one might study of supersonic fluids in order to study black holes or the Unruh effect. In particular, some have suggested this kind of analogy might be used to experimentally probe Hawking radiation.

The supersonic model was first introduced by Unruh (1981), together with an argument for the experimental confirmation of Hawking radiation in this way, and are now commonly referred to as 'acoustic black holes'[1]. The idea, roughly speaking,

[1]See Unruh and Schützhold (2007); philosophers have recently picked up this topic; cf. Dardashti et al. (forthcoming).

is that a 'sound horizon' can be created by accelerating a fluid to speeds faster than the speed of sound waves in the fluid. A small fish passing across this horizon might yell '*Helllppp—*' but no fish outside the horizon would ever hear the sound waves once the small fish crosses the horizon. Moreover, as a small fish approaches the sound horizon, the sound waves would shift to ever-lower frequencies, much like light continuously emitted from a source approaching a black hole horizon.

To make this story more robust, the fluid dynamical model must be describable as a manifold with a Lorentz-signature metric $(\mathbb{R}^4, g_{\mu\nu})$, *not* to represent the metrical structure of spacetime, but to represent the properties of an earthly fluid. The motion of a sound wave through that fluid should then be described by a wave equation on this Lorentzian manifold.

The aim of these notes is to introduce this representation, and in particular the idealisations that must be assumed to hold in order for it to be accurate.

2. Fluid mechanics

Acoustic black holes begin with an entirely standard model of acoustics in fluid mechanics, of the kind that one finds in Landau and Lifshitz (1987, Chapters 1 and 8). I will introduce this model in a way that highlights the physical assumptions underpinning it: the fluid is assumed to be at least approximately without total flux, satisfying Newton's equation with only pressure and gravity as forces, as well as homogeneous, irrotational, bounded, and barotropic.

Fluid dynamics is a branch of classical continuum mechanics. A fluid like air or liquid water is modelled in terms of two smooth scalar fields on a flat Newtonian (or Minkowski[2]) spacetime: a scalar field $\rho : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$, representing the time-dependent density of the fluid in space, and a vector field $\mathbf{v} : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ representing the time-dependent velocity at each spatial point. I will write their values at a point $(\mathbf{x}, t)$ in Euclidean coordinates as $\rho(\mathbf{x}, t)$ and $\mathbf{v}(\mathbf{x}, t)$, respectively. The fundamental assumption of this theory is that, in the absence of sources and sinks, there is no total flux of fluid going in or out of the region. Locally, this is

[2]All velocities will be so small compared to the speed of light that the difference here is negligible.

expressed by the statement that the liquid has zero divergence, through so-called *continuity equation* (I write ∂_t as shorthand for $\frac{\partial}{\partial t}$):

Assumption 1 (flux-free)**.** $\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0$.

The second assumption is that the fluid is homogeneous, in that its mass a each point is proportional to the density. This means that, in appropriate units, the value of the density ρ is the same as that of the mass m, expressed as another scalar field $m(\mathbf{x}, t)$:

Assumption 2 (homogeneity)**.** $\rho(\mathbf{x}, t) = m(\mathbf{x}, t)$ at every point $(\mathbf{x}, t)$.

The third assumption is that Newton's second law is satisfied, $F = ma$, where $F(\mathbf{x}, \mathbf{v}, t)$ is a smooth vector field representing the total force. We can write it using our last assumption that $m = \rho$, and using an expression of acceleration $a = d\mathbf{v}/dt$ for which[3] the time and space derivatives are separated, $d\mathbf{v}/dt = \partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v}$. The result is known as Euler's equation:

Assumption 3 (Euler's equation)**.** $F = ma = \rho\,(\partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v})$.

We assume the presence of only two forces: the uniform force of gravity near the surface of the Earth, and a non-zero force of fluid pressure. Both are assumed to be conservative, in that they can be written as the gradient of a scalar potential: for all $\mathbf{x}, t$,

Assumption 4 (Pressure-Gravity Forces)**.** $F(\mathbf{x}, t) = -\underbrace{\nabla p(\mathbf{x}, t)}_{\text{pressure}} - \underbrace{\rho(\mathbf{x}, t) \nabla \alpha(\mathbf{x})}_{\textit{gravity}}$.

This force is not guaranteed to be time reversal invariant. Although the gravitational term is time-independent, the pressure term is not, and therefore need not satisfy $p(\mathbf{x}, t) = p(\mathbf{x}, -t)$ for all $(\mathbf{x}, t)$ as time reversal invariance requires.[4] This is in particular the case in the presence of non-trivial viscosity, which arises due to the flow

[3] Explicitly: $\frac{d\mathbf{v}}{dt} = \frac{\partial t}{\partial t}\frac{\partial \mathbf{v}}{\partial t} + \frac{\partial x}{\partial t}\frac{\partial \mathbf{v}}{\partial x} + \cdots + \frac{\partial z}{\partial t}\frac{\partial \mathbf{v}}{\partial z} = \frac{\partial \mathbf{v}}{\partial t} + (v_x + v_y + v_z) \cdot (\frac{\partial \mathbf{v}}{\partial x} + \cdots + \frac{\partial \mathbf{v}}{\partial z}) = \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot (\nabla \mathbf{v})$.

[4] Time reversal invariance means that for all $\mathbf{x}$, if $m(d^2/dt^2)\mathbf{x}(t) = \mathbf{F}(\mathbf{x}, t)$ for all t, then $m(d^2/dt^2)\mathbf{x}(-t) = \mathbf{F}(\mathbf{x}, t)$ for all t; equivalently, $\mathbf{F}(\mathbf{x}, t) = \mathbf{F}(\mathbf{x}, -t)$ for all $\mathbf{x}, t$ (in some coordinate system). This holds above if and only if $p(\mathbf{x}, t) = p(\mathbf{x}, -t)$.

of energy into intermolecular degrees of freedom. Some fluid dynamical expressions of black hole radiation rely on time reversal symmetry, for example by running an experimental analogue of a white hole, and concluding the presence of black hole radiation by time reversal symmetry. Such arguments break down in realistic systems with any viscosity. But let us set this concern aside for now: we will instead assume that velocity can be ignored in our 'barotropic' assumption below, which restores time reversal symmetry.

We can now insert our pressure-gravity force into Euler's equation and simplify,[5] which produces the stress-free Navier-Stokes equation:

$$\partial_t \mathbf{v} = -\frac{1}{\rho}\nabla p - \nabla\left(\tfrac{1}{2}\mathbf{v}\cdot\mathbf{v} + \alpha\right) + \mathbf{v}\times(\nabla\times\mathbf{v}). \tag{1}$$

Next we assume the fluid is both irrotational ($\nabla\times\mathbf{v} = 0$), and that the fluid is always contained in a compact (bounded) region of space, with vanishing velocity outside that region:

Assumption 5 (no-vorticity and bounded)**.** $\nabla\times\mathbf{v} = \mathbf{0}$ everywhere, and $\mathbf{v} = 0$ outside a compact subset $S \subset \mathbb{R}^3$, for all times $t \in \mathbb{R}$.

The first (irrotational) property applies quite broadly to typical fluids due to Kelvin's circulation theorem, which states that an irrotational fluid experiencing typical forces (e.g. viscous or other non-conservative stresses) will remain irrotational for all time (c.f. Landau and Lifshitz 1987, §§8-9). So, if the initial conditions of an experiment are created in a laboratory by arranging a fluid at rest and then applying typical forces, the result will be an irrotational fluid. This kills the rotational term ($\nabla\times\mathbf{v}$) in Equation (1).

The second (bounded) property is also reasonable, assuming our fluid lives in a container, though (notably) this assumption is completely unjustified in the analogous context of a black hole spacetime, which is in general spatially unbounded.

[5]Euler's equation becomes, $-\nabla p - \rho\nabla\alpha = \rho(\partial_t\mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{v})$. Rearranging and dividing through by $\rho \neq 0$ gives, $\partial_t\mathbf{v} = -(1/\rho)\nabla p - \nabla\alpha - \mathbf{v}\cdot\nabla\mathbf{v}$, and adding and subtracting $\frac{1}{2}\nabla(\mathbf{v}\cdot\mathbf{v})$ gives $\partial_t\mathbf{v} = -(1/\rho)\nabla p - \nabla\left(\frac{1}{2}\mathbf{v}\cdot\mathbf{v} + \alpha\right) + \frac{1}{2}\nabla(\mathbf{v}\cdot\mathbf{v}) - \mathbf{v}\cdot\nabla\mathbf{v}$. The identity $\frac{1}{2}\nabla(\mathbf{v}\cdot\mathbf{v}) = \mathbf{v}\times(\nabla\times\mathbf{v}) + \mathbf{v}\cdot\nabla\mathbf{v}$ then implies our result.

But *given the spatial boundedness assumption*, we can make use of the Helmholtz-Hodge decomposition theorem, which says under these conditions, $\mathbf{v} = \nabla\Phi + \mathbf{R}$, for some time-dependent scalar field Φ, and for some 'rotational' vector field $\mathbf{R}$ that vanishes when $\nabla \times \mathbf{v} = 0$. Combining these assumptions, we have that $\mathbf{v} = \nabla\Phi$. Our expression of the Navier-Stokes equation then becomes,

$$\partial_t(\nabla\Phi) = \nabla\left(\partial_t\Phi\right) = -\frac{1}{\rho}\nabla p - \nabla\left(\tfrac{1}{2}(\nabla\Phi)\cdot(\nabla\Phi) + \alpha\right). \tag{2}$$

Now, we *almost* have a nice wave equation here, except for the difficult non-linear term $(1/\rho)\nabla p$. Our next assumption will make it more tractable. Suppose the density ρ of the fluid depends only on pressure, a property sometimes referred to as 'barotropy':

Assumption 6 (barotropic)**.** There exists a smooth function $f : \mathbb{R} \to \mathbb{R}$ such that $\rho(\mathbf{x}, t) = f\big(p(\mathbf{x}, t)\big)$ for all $\mathbf{x}, t$.

This assumption ignores the intensive thermodynamic variables associated with realistic fluids, such as temperature or chemical potential, and assumes that density is determined entirely by pressure. For example, when a sound wave propagates through a fluid, it has been known since Laplace (1816) that the process is not isothermal: the compression of molecules in the fluid gives rise to a temperature change that effects the elasticity of the fluid. However, insofar as temperature changes of this kind (and others) are all assumed to be small, the property may be approximately true.

Barotropy allows us to apply another standard trick to express the non-linear term $(1/\rho)\nabla p$ in a more convenient form, as the gradient of a scalar potential. The trick is to define that scalar field as,

$$h(\mathbf{x}, t) := \int_0^{p(\mathbf{x},t)} \frac{1}{f(\lambda)} d\lambda.$$

This definition has been designed to entail[6] the lovely simplification, $(1/\rho)\nabla p = \nabla h$. Substituting this into Equation 2 gives,

$$\nabla\left(\partial_t\Phi + h + \tfrac{1}{2}(\nabla\Phi)\cdot(\nabla\Phi) + \alpha\right) = 0. \tag{3}$$

We can now drop the derivative operator ∇, up to addition of a divergence-free scalar of integration χ (possibly time-dependent), $\nabla\chi = 0$. This is just gauge-fixing: χ can always can always be absorbed into the velocity potential by defining $\Phi \mapsto \Phi' = \Phi + \int_0^t \chi dt$. The integral is divergence-free, so Φ' defines the same physical velocity field, $\mathbf{v} = \nabla\Phi'$. This means that there no new physical assumption introduced by just setting $\chi = 0$ and writing our equation of motion as,

$$\boxed{\partial_t\Phi + \tfrac{1}{2}(\nabla\Phi)\cdot(\nabla\Phi) + h + \alpha = 0.} \tag{4}$$

Let me summarise this section. A fluid mechanical system consists in a triple $(\mathbf{v}, \rho, p)$. Suppose this triple has the properties assumed above: it is flux-free, homogeneous, bounded, irrotational, barotropic, satisfies Newton's equation, and has only the conservative forces of pressure and gravitation. Then scalar fields (Φ, h) exist such that $\mathbf{v} = \nabla\Phi$ and $\nabla h = (1/\rho)\nabla p$, and such that the motion of the system is characterised by Equation (4). However, it is also worth emphasising the ways our derivation of this wave equation can break down:

- The fluid might have non-trivial inhomogeneities
- One might not be able to ignore thermodynamic properties, like viscosity or temperature in the fluid;
- The fluid might not be assumed or vorticity free.

3. Sound Waves

In this section we construct the perturbation model of sound waves in the fluid and, following Unruh (1981), show that they exhibit a Lorentzian geometry.

[6]The calculation is a simple application of the fundamental theorem of calculus: our definitions imply $\nabla h = (\nabla p)(\partial/\partial p)h = (\nabla p)(\partial/\partial p)\int_0^p d\lambda/f(\lambda)$, which by the fundamental theorem is equal to $(\nabla p)(1/f(p))$, which is just $(\nabla p)/\rho$.

A sound wave in a continuous fluid consists successive oscillations in pressure passing across it.[7] It's hard to get an exact solution to this complex behaviour satisfying Equation (4). However, we can still give an accurate approximation of a sound wave solution as a linear perturbation of a fluid without sound waves. To do this, we first consider a velocity potential Φ_0 and pressure-density potential h_0 that solve Equation (4) in the absence of sound waves. Call these the 'equilibrium' potentials. The potentials corresponding to an ordinary sound wave are only small displacements of these potentials, since the pressure oscillations are so small. As a physical assumption, this says:

Assumption 7 (sound-perturbation)**.** If a solution (Φ_0, h_0) to the wave equation (4) represents a fluid with no sound waves, then the same fluid with sound waves can be represented by a solution (Φ, h), related to the first by functions $\Phi = \Phi_0 + \phi$ with $|\phi| << |\Phi_0|$ and $h = h_0 + \psi$ with $|\psi| << |h_0|$ (assumed to be analytic).

Analyticity is reasonable because of our assumption that the fluid is bounded in a container: then the Stone-Weierstrass theorem says every continuous function has an analytic approximation. Moreover, the fluctuations ϕ, ψ associated with sound wave potentials really are extremely small compared to the equilibrium potentials, at least for typical fluids. Unfortunately, this approximation fails exactly in the case of interest, when the fluid is accelerated to speeds at which large shock waves develop. This usually happens already when the fluid moves at about 1/3 the speed of sound (Landau and Lifshitz 1987, §122). Unruh (1981) suggests that this instability in the model might be avoided through clever experimental techniques. Not one to underestimate the cleverness of experimentalists, I will follow his lead.

We can use perturbation theory to give a first-order (linear) approximation of the sound wave motion, and it will be accurate to the extent of our assumption that $|\phi|$ and $|\psi|$ are small. Writing Equation (4) as $F(\Phi, h) = 0$ with F analytic, we consider its expression for our sound wave solution (Φ, h) in the two-variable

[7] These oscillations affect density and temperature as well; however, we have already assumed in the discussion above that these can be ignored.

expansion around (Φ_0, h_0):

$$F(\Phi, h) = F(\Phi_0, h_0) + \phi \frac{\partial}{\partial \Phi} F\big|_{(\Phi_0,h_0)} + \psi \frac{\partial}{\partial h} F\big|_{(\Phi_0,h_0)} + (\text{h.o. terms}). \tag{5}$$

We can now drop the higher-order (h.o.) terms and calculate $(\partial/\partial\Phi)F$ and $(\partial/\partial h)F$ explicitly,[8] which gives us our accurate linear approximation of the statement that $0 = F(\Phi, h)$. Writing $\mathbf{v}_0 = \nabla\Phi_0$,

$$0 = F(\Phi, h) \approx \frac{1}{\rho} \left(\partial_t + \mathbf{v}_0 \cdot \nabla + \nabla \cdot (\nabla\Phi_0)\right) \frac{\rho}{c^2} \left(\partial_t + \mathbf{v}_0 \cdot \nabla\right) \phi - h_0 \nabla\phi, \tag{6}$$

where the speed of the adiabatic sound wave is defined to be $c^2 = dp/d\rho$, and use the fact that $dp/d\rho = (dp/dh)(dh/d\rho) = \rho(dh/d\rho)$ by the definition of h, and hence $c^2 = \rho(dh/dp)$. When our assumptions are satisfied, Equation (6) approximates the propagation of a sound wave in the fluid. It is sometimes called the 'linearised wave equation' for a sound wave.

With Equation (6), we have arrived at our equation of motion for a fluid containing a sound wave travelling at speed c. It is an approximation, but a good one given our assumptions. The remainder of this section will be devoted to expressing this equation of motion in a particularly compact form, which also clarifies its relationship to Lorentzian geometry.

To begin, we can vastly simplify Equation (6) by defining the matrix,

$$\tilde{g}^{\mu\nu} := \frac{\rho_0}{c^2} \begin{pmatrix} 1 & v_0^x & v_0^y & v_0^z \\ v_0^x & (v_0^x v_0^x - c^2) & v_0^x v_0^y & v_0^x v_0^z \\ v_0^y & v_0^y v_0^x & (v_0^y v_0^y - c^2) & v_0^y v_0^z \\ v_0^z & v_0^z v_0^x & v_0^z v_0^y & (v_0^z v_0^z - c^2) \end{pmatrix}. \tag{7}$$

This matrix has Lorentz signature, and through explicit matrix multiplication, we find that the linearised wave equation above becomes simply, $\partial_\mu \tilde{g}^{\mu\nu} \partial_\nu \phi = 0$. We can make that even more familiar by multiplying it by a constant $1/\sqrt{-\tilde{g}}$, where $\tilde{g} = \det \tilde{g}^{\mu\nu}$, and defining,

$$g^{\mu\nu} := \frac{1}{\sqrt{-\tilde{g}}} \tilde{g}^{\mu\nu}. \tag{8}$$

[8]Calculation didn't fit in the margins.

Then our linearised wave equation becomes the massless Klein-Gordon equation[9] for a Lorentzian manifold $(\mathbb{R}^4, g_{\mu\nu})$,

$$\partial_\mu \left(\sqrt{-g}\, g^{\mu\nu}\partial_\nu\psi\right) = 0. \tag{9}$$

Defining $d\tau = (dt + \frac{v^i}{(c^2-v^2)}dx^i)$, which is always possible when $\frac{v^i}{(c^2-v^2)}$ is an integrable vector field, and assuming radial symmetry, Unruh (1995) finds a line element that looks just like the Schwarzschild black hole,

$$ds^2 = \rho\left((c^2 - v^2)d\tau^2 - \frac{1}{1-\frac{v^2}{c^2}}dr^2 - r^2 d\Omega^2\right), \tag{10}$$

where $d\Omega^2 = d\theta^2 + sin^2\theta d\phi^2$. Of course, it has been derived here, nothing in the equation of motion represents a massless relativistic field in spacetime; it is just the (approximate) equation of motion for a fluid containing a sound wave. However, putting it in this way brings out a number of interesting features of that equation:

- Informally: the fluid satisfies $v_0^2 < c^2$ in some regions and $v_0^2 > c^2$ in others with respect to the speed of sound c, then it admits a certain kind of 'horizon' beyond which sound waves cannot return.
- This equation of motion exhibits an interesting symmetry: it is preserved along the 'static' vector field $(\partial/\partial t)^\mu = (1, 0, 0, 0)$, since the tensor $g_{\mu\nu}$ is preserved there. Formally, this is equivalent to the vector field being a *Killing field.*
- Moreover, the Killing field satisfies $(\partial/\partial t)^\mu(\partial/\partial t)_\mu = -(c^2 - v_0^2)$, and so vanishes on the 'null' surfaces $\pm v_0 = c$; each of these surfaces is called a *Killing horizon.*
- These horizons intersect at a two-dimensional surface S, which makes them into what is known as a *bifurcate Killing horizon.*

As we have seen in other discussions, these are exactly the conditions under which we may discuss the generalised Unruh effect. However, it is less clear whether an

[9]This is the equation of motion for a free massless test field ϕ, $\Box^2\phi = \nabla^\mu\nabla_\mu\phi = g^{\mu\nu}\nabla_\mu\nabla_\nu\phi = 0$, where ∇_μ is the unique covariant derivative compatible with $g_{\mu\nu}$. For a scalar field, the covariant and coordinate derivatives are the same ($\nabla_\mu\phi = \partial_\mu\phi$), and so we get Equation (8).

analogue of Hawking radiation is possible, since the manifold on which this metric is defined is assumed to have bounded volume.

References

Dardashti, R., Thébault, K. P. Y. and Winsberg, E. (forthcoming). Confirmation via Analogue Simulation: What Dumb Holes Could Tell Us about Gravity, *British Journal for the Philosophy of Science* .

Landau, L. D. and Lifshitz, E. M. (1987). *Fluid Mechanics*, Vol. 6 of *Course of Theoretical Physics*, 2nd edn, Pergamon Press. Translated from Russian by J.B. Sykes and W.H. Reid.

Laplace, P. S. (1816). Sur la vitesse du son dans l'air et dans l'eau, *Annales de Chimie et de Physique* **3**: 238–241.

Unruh, W. G. (1981). Experimental black-hole evaporation?, *Physical Review Letters* **46**(21): 1351–3.

Unruh, W. G. and Schützhold, R. (eds) (2007). *Quantum Analogues: From Phase Transitions to Black Holes and Cosmology*, Springer-Verlag Berlin Heidelberg.