The Astrophysical Journal, 959:34 (25pp), 2023 December 10 https://doi.org/10.3847/1538-4357/acf659
© 2023. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

Monster Radiative Shocks in the Perturbed Magnetospheres of Neutron Stars

Andrei M. Beloborodov[1,2]
[1] Physics Department and Columbia Astrophysics Laboratory, Columbia University, 538 West 120th Street, New York, NY 10027, USA; amb@phys.columbia.edu
[2] Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany
Received 2022 October 24; revised 2023 August 31; accepted 2023 September 1; published 2023 December 1

Abstract

Magnetospheres of neutron stars can be perturbed by star quakes, interaction in a binary system, or sudden collapse of the star. The perturbations are typically in the kilohertz band and excite magnetohydrodynamic waves. We show that compressive magnetospheric waves steepen into monster shocks, possibly the strongest shocks in the Universe. The shocks are radiative, i.e., the plasma energy is radiated before it crosses the shock. As the kilohertz wave with the radiative shock expands through the magnetosphere, it produces a bright X-ray burst. Then, it launches an approximately adiabatic blast wave, which will expand far from the neutron star. These results suggest a new mechanism for X-ray bursts from magnetars and support the connection of magnetar X-ray activity with fast radio bursts. Similar shocks may occur in magnetized neutron-star binaries before they merge, generating an X-ray precursor of the merger. Powerful radiative shocks are also predicted in the magnetosphere of a neutron star when it collapses into a black hole, producing a bright X-ray transient.

Unified Astronomy Thesaurus concepts: X-ray transient sources (1852); Neutron stars (1108); Magnetars (992); Radiative processes (2055); Radio bursts (1339); Plasma astrophysics (1261)

1. Introduction

Magnetospheres of neutron stars are formed by plasma immersed in a strong magnetic field $B_{\rm bg}$. They have a high magnetization parameter,

$$\sigma_{\rm bg} \equiv \frac{B_{\rm bg}^2}{4\pi\rho_{\rm bg}c^2} \gg 1, \tag{1}$$

where $\rho_{\rm bg}$ is the plasma mass density and c is the speed of light; subscript "bg" stands for "background" for waves investigated in this paper. The closed magnetosphere is approximately dipole at radii r much greater than the star radius $R_\star$. It ends and a magnetized wind begins near the light cylinder $R_{\rm LC} = c/\Omega$, where Ω is the star rotation rate. This basic picture is confirmed by extensive studies of pulsars (Philippov & Kramer 2022).

1.1. Perturbations of Neutron Star Magnetospheres

In some cases, the magnetosphere becomes significantly perturbed. In particular, quakes in magnetars launch magnetospheric waves (Blaes et al. 1989; Thompson & Duncan 1996; Bransgrove et al. 2020). Some mechanism quickly dissipates the waves, generating X-ray bursts that are observed as the main form of magnetar activity (Kaspi & Beloborodov 2017). Significant perturbations are also expected in tight neutron-star binaries, where the two magnetospheres interact with each other. A huge disturbance of a neutron star magnetosphere occurs when the star collapses into a black hole (Lehner et al. 2012). In all of these cases, the excited waves are typically in the kilohertz band.

Such perturbations are well described by magnetohydrodynamics (MHD), which supports waves of the perturbed magnetic field and electric field $\boldsymbol{E}\perp\boldsymbol{B}_{\rm bg}$. These MHD modes can have wavevectors $\boldsymbol{k}$ in any direction and can be of two types: (1) shear Alfvén waves with $\boldsymbol{k}\cdot\boldsymbol{E}\neq 0$, and (2) compressive waves with $\boldsymbol{E}\perp\boldsymbol{k}$, so-called "fast magnetosonic modes." Both modes propagate with an ultrarelativistic group speed,

$$\frac{v_{\rm gr}}{c} \approx 1 - \sigma_{\rm bg}^{-1}. \tag{2}$$

Phase speed $v_{\rm ph} = \omega/k$ of the magnetosonic waves is also close to c while Alfvén waves have $v_{\rm ph} \approx c\cos\alpha$, where α is the angle between $\boldsymbol{k}$ and $\boldsymbol{B}_{\rm bg}$. A detailed discussion of waves in $e^\pm$ plasma in a strong background magnetic field $B_{\rm bg}$, including two-fluid and single-fluid (MHD) descriptions, is found in Arons & Barnard (1986).

1.2. MHD Waves from Quakes

Substantial work was devoted to the relativistic Alfvén waves excited by magnetar quakes (e.g., Thompson & Blaes 1998; Troischt & Thompson 2004; Yuan et al. 2020; Nättilä & Beloborodov 2022), and little attention was given to the magnetosonic waves. In fact, quakes can excite both modes.

Neutron star quakes involve shear oscillations of the crust with horizontal displacements $\delta\boldsymbol{r}$. Shear waves inside the neutron star crust propagate with speed $v_{\rm sh} = \omega/k_{\rm sh} \approx 10^8\,{\rm cm\,s^{-1}}$ (Blaes et al. 1989). Their characteristic lowest frequency is $\omega \sim v_{\rm sh}/h \sim 10^4\ {\rm rad\,s^{-1}}$, where h is the hydrostatic scale height. A large fraction of the quake energy may be at much higher frequencies, e.g., $\omega \sim 10^5\,{\rm rad\,s^{-1}}$. The crustal waves leak to the magnetosphere with a transmission coefficient $\mathcal{T} \sim 0.1\,\omega_5^{3/5}B_{15}^{2/5}$ (Blaes et al. 1989; Bransgrove et al. 2020).

To quickly see that quakes can emit both Alfvén and magnetosonic modes, one can consider crustal oscillations with $\omega \gg c/R_\star$ and a wavevector that is exactly radial in a region of the stellar surface (Figure 1). Such oscillations satisfy $\partial_\phi = \partial_\theta = 0$ in spherical coordinates r, θ, ϕ, and this symmetry is preserved during wave transmission, so the emitted

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

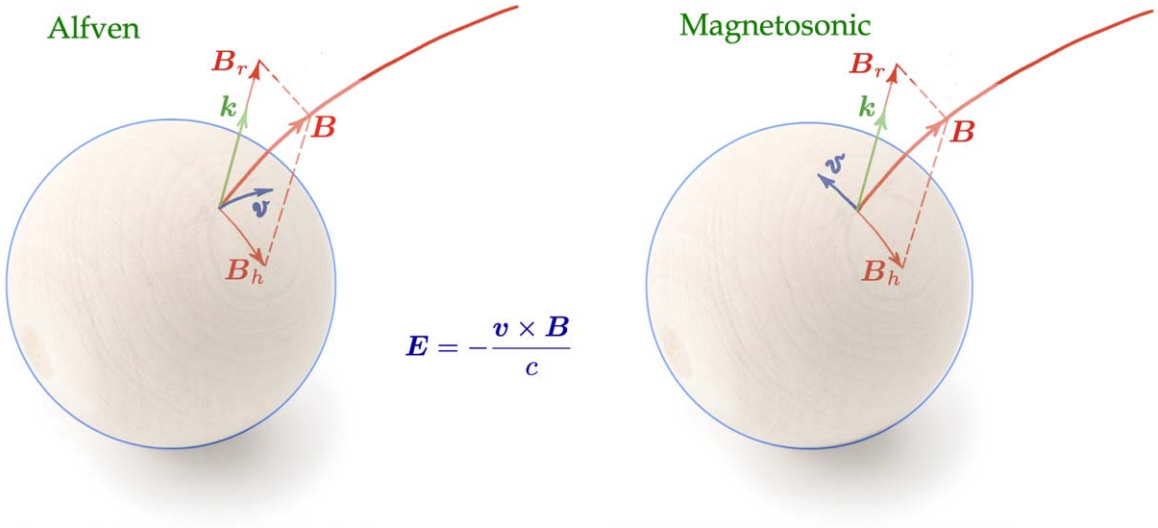

Figure 1. Excitation of magnetospheric waves by a star quake with $\omega \gg c/R_\star$ and a vertical (radial) wavevector $\boldsymbol{k}$. The wave polarization is set by the direction of the electric field $\boldsymbol{E} = \boldsymbol{B} \times \boldsymbol{v}/c$ in the crustal shear oscillation with a horizontal velocity $\boldsymbol{v}$. Left: crustal motion with $\boldsymbol{v} \| \boldsymbol{n} \equiv \boldsymbol{k} \times \boldsymbol{B}_{\rm bg}$ excites an Alfvén wave $\boldsymbol{E} \perp \boldsymbol{n}$. Right: $\boldsymbol{v} \perp \boldsymbol{n}$ excites a magnetosonic wave $\boldsymbol{E} \| \boldsymbol{n}$. The wave amplitude is small, so $\boldsymbol{B} \approx \boldsymbol{B}_{\rm bg}$ near the star.

magnetospheric wave also has a radial $\boldsymbol{k}$. This simple setup is reduced to a locally plane-parallel transmission problem; it was used by Blaes et al. (1989) to study how quakes with $\delta \boldsymbol{r} \| (\boldsymbol{k} \times \boldsymbol{B}_{\rm bg})$ excite Alfvén waves in the magnetosphere. It is easy to see that quakes with $\delta \boldsymbol{r} \nparallel (\boldsymbol{k} \times \boldsymbol{B}_{\rm bg})$ emit magnetosonic modes. Polarization of the emitted wave is set by the electric field in the crustal oscillation, $\boldsymbol{E} = \boldsymbol{B}_{\rm bg} \times \boldsymbol{v}/c$ (where $\boldsymbol{v} = -i\omega \delta \boldsymbol{r}$), and its type is determined by the orientation of $\boldsymbol{E}$ relative to the vector $\boldsymbol{n} \equiv \boldsymbol{k} \times \boldsymbol{B}_{\rm bg}$. It is a pure Alfvén mode if $\boldsymbol{E} \perp \boldsymbol{n}$ and a pure magnetosonic wave if $\boldsymbol{E} \| \boldsymbol{n}$. The condition for magnetosonic excitation $\boldsymbol{E} \cdot \boldsymbol{n} \neq 0$ may be stated as

$$(\boldsymbol{B}_{\rm bg} \times \boldsymbol{v}) \cdot (\boldsymbol{k} \times \boldsymbol{B}_{\rm bg}) = (\boldsymbol{B}_{\rm bg} \cdot \boldsymbol{v})(\boldsymbol{B}_{\rm bg} \cdot \boldsymbol{k}) \neq 0. \quad (3)$$

The transmitted quake power is partitioned between the Alfvén and magnetosonic modes as $E_{\rm A}^2/E_{\rm ms}^2 = \tan^2\psi$ where ψ is the angle between the quake motion $\boldsymbol{v}$ and the horizontal component of the magnetic field $\boldsymbol{B}_{\rm bg}^{\rm h}$.

A complete picture of wave emission is complicated by several additional effects:

(1) In addition to the vertical (radial) wavevector k_r, crustal waves have a horizontal component $k_{\rm h}$ (Bransgrove et al. 2020) and excite nonradial magnetospheric waves. As an example, Yuan et al. (2020) examined wave generation in a uniform oblique $\boldsymbol{B}_{\rm bg}$ attached to a horizontal conducting surface perturbed by axisymmetric horizontal shear. They showed that the shear excites a mixture of Alfvén and magnetosonic waves. These calculations should be extended in the future to include a realistic density profile of the crust.

(2) Waves of lower frequencies $\omega \lesssim c/R_\star$ develop at larger radii $R_{\rm w} \sim R_\star + c/\omega$. In the region $R_\star < r < R_{\rm w}$, the magnetosphere adjusts to the surface oscillation in a quasi-static manner and, effectively, the magnetospheric footpoints in the wave are relocated from the stellar surface to the sphere of radius $r \approx R_{\rm w}$. The quasi-static deformation of the magnetosphere at $r \lesssim R_{\rm w}$ differs from the crustal deformation and needs further investigation of its compressive component that could drive a small-amplitude magnetosonic wave. For example, a strongly twisted magnetosphere in axisymmetric equilibrium responds to additional surface shear $\partial_\theta v_\phi \neq 0$ by inflating at $r \lesssim R_{\rm w}$ (Parfrey et al. 2013).

(3) The twisted magnetospheres near magnetars have significant spatial variations of the toroidal magnetic field component B_ϕ, so vector $\boldsymbol{n} = \boldsymbol{k} \times \boldsymbol{B}_{\rm bg}$ can change its direction on a scale comparable to r. Then, any attempt to launch a pure Alfvén mode at $r \sim R_{\rm w} \sim c/\omega$ (for $\omega \lesssim c/R_\star$) will inevitably generate a mixture of the Alfvén and magnetosonic waves because the two linear modes have different refraction indices $c/v_{\rm ph}$ and will be unable to adiabatically track the changing local $\boldsymbol{n}$ until propagating to $r \gg c/\omega$.

Pure Alfvénic excitations occur in the simple case of axisymmetric quakes with azimuthal $\boldsymbol{v}$ and an untwisted background magnetosphere (then $k_\phi = 0$ and $B_{\rm bg}^\phi = 0$, so $\boldsymbol{n} \| \boldsymbol{v}$). Even in this case, the emitted Alfvén waves can convert to magnetosonic waves, because of nonlinear effects. Alfvén waves have $\boldsymbol{v}_{\rm gr} \| \boldsymbol{B}_{\rm bg}$, so they are ducted along the closed magnetic loops, and the nonlinear conversion peaks when the wave reaches the top of the loop (Yuan et al. 2021). In addition, at later times, the bouncing Alfvén waves trapped in the loop develop a nonlinear turbulent cascade, which emits magnetosonic waves (e.g., Li et al. 2019).

In addition to sudden quakes, the crust may flow plastically and slowly twist a magnetospheric flux bundle to an instability threshold. Three-dimensional simulations of this process demonstrate that relaxation of the unstable flux bundle generates fast magnetosonic waves in the magnetosphere (Mahlmann et al. 2023).

1.3. Shocks

In the limit of high $\sigma_{\rm bg}$, the emitted magnetosonic waves are equivalent to electromagnetic radio waves propagating without coupling to the magnetized plasma: the oscillating $\boldsymbol{E} \perp \boldsymbol{B}_{\rm bg}$ drives a tiny electric current, negligible compared to the displacement current $\partial_t \boldsymbol{E}/4\pi$. Therefore, compressive perturbations of the magnetosphere are usually assumed to propagate as vacuum electromagnetic waves superimposed linearly with

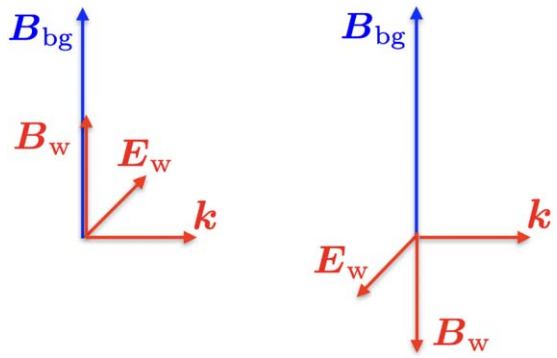

Figure 2. A magnetosonic wave with wavevector $\boldsymbol{k}\perp\boldsymbol{B}_{bg}$ at two oscillation phases: when the wave field $\boldsymbol{B}_w=\boldsymbol{B}-\boldsymbol{B}_{bg}$ is aligned with $\boldsymbol{B}_{bg}$ (left) and antialigned with $\boldsymbol{B}_{bg}$ (right). The wave electric field $\boldsymbol{E}_w=\boldsymbol{E}$ is always orthogonal to $\boldsymbol{B}_{bg}$. The quantity B^2-E^2 reaches zero when $\boldsymbol{B}_w=-\boldsymbol{B}_{bg}/2$. At this moment, the plasma drift speed $\boldsymbol{v}=c\,\boldsymbol{E}\times\boldsymbol{B}/B^2$ approaches c, so it experiences ultrarelativistic acceleration in the direction opposite to the wave propagation direction.

the background field $\boldsymbol{B}_{bg}$, and freely escape.[3] The linear propagation would imply no shock formation.

However, this simple picture is safe only for low-amplitude waves, $|\boldsymbol{E}|\ll B_{bg}$, i.e., near the neutron star where B_{bg} is strong. As the wave expands to larger radii r, the dipole background field decreases as $B_{bg}\propto r^{-3}$ while the wave amplitude E_0 decreases as r^{-1}. Their ratio E_0/B_{bg} grows as r^2 and eventually approaches unity. It is easy to see that the linear propagation of an oscillating wave becomes impossible when

$$2E_0\sin\alpha > B_{bg}, \tag{4}$$

where α is the angle between $\boldsymbol{B}_{bg}$ and the wavevector $\boldsymbol{k}$. Indeed, the linear superposition of $\boldsymbol{B}_{bg}$ with the vacuum electromagnetic wave ($\boldsymbol{E}_w=\boldsymbol{E}$ and $\boldsymbol{B}_w=\boldsymbol{B}-\boldsymbol{B}_{bg}$) gives

$$B^2-E^2=B_{bg}^2+2\boldsymbol{B}_{bg}\cdot\boldsymbol{B}_w, \tag{5}$$

where we used $\boldsymbol{E}_w^2=\boldsymbol{B}_w^2$. MHD description breaks if $B^2<E^2$. This occurs when the condition in Equation (4) is met.

For instance, consider the equatorial plane of a dipole magnetosphere. Here, a spherical wave front propagates with $\boldsymbol{k}\perp\boldsymbol{B}_{bg}$ ($\alpha=\pi/2$) and the magnetic field perturbation $\boldsymbol{B}_w$ oscillates along $\boldsymbol{B}_{bg}$ (Figure 2). The linear superposition wave +background gives minimum B^2-E^2 at $B_{min}=B_{bg}-E_0$, reaching $B^2-E^2=0$ when $E_0=B_{bg}/2$. Note that the plasma oscillates with the drift speed $\boldsymbol{v}/c=\boldsymbol{E}\times\boldsymbol{B}/B^2$, and $E^2\to B^2$ corresponds to $|\boldsymbol{v}|\to c$. This implies a runaway growth of the plasma kinetic energy, and the MHD wave can no longer be approximated as a vacuum electromagnetic wave. Such energy conversion is a general effect of $E^2\to B^2$ (it also happens in Alfvén waves; Li et al. 2019, 2021; see also Levinson 2022 for a recent discussion). In magnetosonic waves, energy conversion prevents reaching $E^2=B^2$ by steepening the wave into a shock. The appearance of shocks at $E_0=B_{bg}/2$ was previously noted in the context of waves in pulsar winds (Lyubarsky 2003). A fully kinetic (local-box) particle-in-cell simulation of a magnetosonic wave propagating through a decreasing B_{bg} has recently been performed by Chen et al. (2022). Their simulation demonstrates the sudden steepening of the wave into a shock when E^2 approaches B^2.

Waves consisting of half oscillation with $\boldsymbol{B}_w\cdot\boldsymbol{B}_{bg}>0$ and no part with $\boldsymbol{B}_w\cdot\boldsymbol{B}_{bg}<0$ would never face the $E^2=B^2$ limit. One may think that in this case the wave avoids shock formation. However, as explained below, such half-waves also steepen into shocks, although this occurs gradually, at a larger radius. This gradual steepening creates a forward shock, launching an accelerating blast wave that expands far beyond the magnetosphere.

1.4. This Paper

Since shocks appear to be a generic outcome of magnetosonic perturbations, a few questions arise: How strong are the shocks? What fraction of the original wave energy gets dissipated by the shock? What fraction of the dissipated energy is radiated?

These questions can be answered by solving a well-defined MHD problem with a simple initial condition: launch a spherically expanding magnetosonic wave with an initial amplitude $E_0/B_{bg}\ll 1$ so that initially (at small radii) it behaves as a vacuum electromagnetic wave. Then, track its expansion through the dipole magnetosphere and examine how it steepens into a shock and propagates afterward. This problem is solved in the present paper. We will show that the plasma Lorentz factor γ in the magnetospheric shocks caused by $E^2\to B^2$ reaches huge values $\gamma\propto\sigma_{bg}$, likely making them the strongest shocks in the Universe. Therefore, we call them "monster shocks." They differ from normal collisionless plasma shocks because they are highly radiative: we will show that the plasma approaching the monster shock radiates its kinetic energy before forming the downstream flow.

Tracking the evolution of magnetospheric waves with shock formation presents an interesting technical challenge. The magnetization parameter σ_{bg} at radii of main interest can exceed 10^{10} (Section 2.1). The large σ_{bg} is usually replaced by $\sigma_{bg}\to\infty$, which corresponds to taking the force-free electrodynamics (FFE) limit of MHD. However, FFE cannot describe shock formation. FFE neglects plasma inertia, so its wave modes propagate with exactly speed of light and cannot steepen.

Since FFE does not provide a suitable framework, one has to examine the problem using the full MHD equations. The equations state conservation of energy and momentum and could be solved numerically with customary discretization methods. However, in practice such methods fail at high magnetization σ (most existing MHD codes have to keep $\sigma<100$ to avoid numerical issues). We take a different approach: we solve the MHD equations along characteristics. It provides an efficient method for both finding and understanding the solution at arbitrarily high σ_{bg}. The solution is easiest to find for short waves, with wavelength $\lambda\ll r$.

Calculations in this paper are performed for axisymmetric wave packets ($\partial_\phi=0$). Setting $\partial_\phi=0$ should also be a good approximation more generally for waves far from their source, where wavevectors $\boldsymbol{k}$ are radial. The wave evolution along the radial ray is controlled by the local $\boldsymbol{B}_{bg}$ encountered along the ray; we are interested in the simple case of a dipole $\boldsymbol{B}_{bg}$.

Formulation of the problem and our method for solving it are described in Section 2. Section 3 explains how we track shocks in MHD waves and then Section 4 presents the full numerical simulation, performed for an equatorial wave in a dipole

[3] It was also proposed that QED effects in ultrastrong B_{bg} could steepen a high-frequency wave into a shock (Heyl & Hernquist 2005). Such effects are absent for kilohertz waves.

magnetosphere. Remarkably, the problem also admits a complete analytical description, which is given in Section 5. Emission from the monster shock is discussed in Section 6. The results are discussed in Section 7.

2. Nonlinear Magnetosonic Waves

MHD fluid is described by the plasma mass density ρ, velocity $\boldsymbol{v} = c\boldsymbol{\beta}$, magnetic field $\boldsymbol{B}$, and electric field $\boldsymbol{E}$. We wish to investigate magnetosonic waves launched in a dipole magnetosphere and find their nonlinear evolution. The term "nonlinear" here means that (a) the wave oscillation is not necessarily small compared to the background field $\boldsymbol{B}_{\rm bg}$, and (b) the wave can be strongly deformed during its evolution and form shocks.

2.1. Unperturbed Background Magnetosphere

The wave will be shown to experience strong evolution at radii r much greater than the neutron star radius $R_\star \sim 10^6$ cm, but well inside the light cylinder $R_{\rm LC}$ (all known magnetars in our galaxy have $R_{\rm LC} \gtrsim 10^{10}$ cm). Our calculations below use the background at radii $R_\star \ll r \ll R_{\rm LC}$, the magnetosphere is approximated as dipole, and its rotation is neglected. Thus, for simplicity, we neglect any twists of the outer magnetosphere, i.e., assume $B_{\rm bg}^{\phi} \approx 0$ at radii of interest. The electromagnetic field of a static dipole magnetosphere is

$$\boldsymbol{E}_{\rm bg} = 0, \qquad \boldsymbol{B}_{\rm bg} = \frac{2\mu\cos\theta}{r^3}\boldsymbol{e}_r + \frac{\mu\sin\theta}{r^3}\boldsymbol{e}_\theta, \tag{6}$$

where $(\boldsymbol{e}_r,\ \boldsymbol{e}_\theta,\ \boldsymbol{e}_\phi)$ is the normalized basis in spherical coordinates r, θ, ϕ with the polar axis along the magnetic dipole moment $\boldsymbol{\mu}$. Magnetars have $\mu \sim 10^{33}$ G cm^3.

Plasma density in the unperturbed static magnetosphere will be approximated by

$$\rho_{\rm bg} = \frac{\mathcal{N}m}{r^3}, \qquad \boldsymbol{v}_{\rm bg} \approx 0, \tag{7}$$

where the dimensionless parameter $\mathcal{N} \equiv n_{\rm bg} r^3$ is approximately constant with r (Beloborodov 2020); m is the particle ($e^\pm$) mass, and $n_{\rm bg} = \rho_{\rm bg}/m$. The lowest $\mathcal{N} \sim 10^{31}\,\mu_{33}(\Omega/{\rm rad\ s^{-1}})$ is set by the Goldreich–Julian density. The more typical $\mathcal{N}$ for magnetars is much higher, $\mathcal{N} \sim 10^{37}$, due to strong electric currents near the star accompanied by $e^\pm$ pair creation with a high multiplicity (Beloborodov 2021a).[4] The magnetization parameter of the background plasma $\sigma_{\rm bg} = B_{\rm bg}^2/4\pi\rho_{\rm bg}c^2$ is

$$\sigma_{\rm bg} \approx \frac{D}{r^3} \sim 10^{10}\mu_{33}^2\mathcal{N}_{37}^{-1}r_8^{-3}, \qquad D \equiv \frac{\mu^2}{4\pi\mathcal{N}mc^2}. \tag{8}$$

The magnetosphere at $r > 10^7$ cm is populated with mildly relativistic $e^\pm$, as they are decelerated by drag exerted by the magnetar radiation (Beloborodov 2013). We neglect these background motions; thus, the plasma energy density (including rest mass) in the unperturbed background ahead of the wave is approximated as $\rho_{\rm bg}c^2$.

[4] The created pairs flow along the magnetic field lines with a speed controlled by the magnetar's radiation field (Beloborodov 2013). In the outer closed magnetosphere, the plasma accumulates and annihilates at $\theta \approx \pi/2$, in a layer of high $e^\pm$ density controlled by annihilation balance and the thickness of the annihilation layer. The $e^\pm$ outflow terminated by the annihilation sink at $\theta = \pi/2$ may have $\mathcal{N} \sim 10^{37}$ at all θ except the thin equatorial layer.

2.2. Axisymmetric Waves

We will examine the radial expansion of an axisymmetric magnetosonic wave. Its electric field $\boldsymbol{E}$ is perpendicular to both $\boldsymbol{B}_{\rm bg}$ and the (radial) propagation direction, so vector $\boldsymbol{E}$ oscillates along ϕ. It is convenient to define scalar E by

$$\boldsymbol{E} = -E\boldsymbol{e}_\phi, \qquad E \equiv -E_\phi. \tag{9}$$

The wave magnetic field is perpendicular to $\boldsymbol{E}$ and oscillates in the poloidal plane,

$$\boldsymbol{B} = B_r\boldsymbol{e}_r + B_\theta\boldsymbol{e}_\theta. \tag{10}$$

Any axisymmetric wave can be described by a toroidal electromagnetic potential,

$$\boldsymbol{A} = A(t, r, \theta)\boldsymbol{e}_\phi, \qquad c\boldsymbol{E} = -\partial_t\boldsymbol{A}, \qquad \boldsymbol{B} = \nabla\times\boldsymbol{A}. \tag{11}$$

This class of fields includes the static dipole background as a special case, with

$$\boldsymbol{A}_{\rm bg}(r, \theta) = \frac{\mu\sin\theta}{r^2}\,\boldsymbol{e}_\phi. \tag{12}$$

Charge density remains zero in magnetosonic waves, $4\pi\rho_e = \nabla\cdot\boldsymbol{E} = 0$. Electric current density $\boldsymbol{j}$ satisfies the Maxwell equation $4\pi\boldsymbol{j} = -\partial_t\boldsymbol{E} + c\,\nabla\times\boldsymbol{B}$. In axisymmetric waves, the current is toroidal:

$$4\pi\boldsymbol{j} = \left\{\partial_t E + \frac{c}{r}[\partial_r(rB_\theta) - \partial_\theta B_r]\right\}\boldsymbol{e}_\phi. \tag{13}$$

Note also that $\boldsymbol{E}$ are $\boldsymbol{B}$ are related by induction equation $\partial_t\boldsymbol{B} = -c\nabla\times\boldsymbol{E}$ (the identity $\partial_t\nabla\times\boldsymbol{A} = \nabla\times\partial_t\boldsymbol{A}$):

$$\partial_t(rB_\theta) = -c\,\partial_r(rE), \qquad \partial_t(rB_r) = \frac{c\,\partial_\theta(E\sin\theta)}{\sin\theta}. \tag{14}$$

As a concrete example, we will consider a wave launched with an initial sine profile

$$E(\xi) = E_0\sin(\omega\xi), \qquad \xi = t - \frac{r}{c}. \tag{15}$$

One full oscillation corresponds to $0 < \xi < \lambda/c = 2\pi/\omega$.

2.3. Short Waves ($\lambda \ll$ r)

The condition $\lambda \ll r$ considerably simplifies the wave propagation problem. In particular, it leads to a simple expression for current $\boldsymbol{j}$, which will be used in Section 2.4 to describe the wave–plasma interaction.

Let us define the wave potential $\boldsymbol{A}_{\rm w} \equiv (A - A_{\rm bg})\boldsymbol{e}_\phi$. It determines the wave fields and $\boldsymbol{j}$:

$$\boldsymbol{B}_{\rm w} \equiv \boldsymbol{B} - \boldsymbol{B}_{\rm bg} = \nabla\times\boldsymbol{A}_{\rm w}, \qquad c\boldsymbol{E} = -\partial_t\boldsymbol{A}_{\rm w}, \tag{16}$$

$$\boldsymbol{j} = [\partial_t(rE) + c\,\partial_r(rB_{\rm w}^\theta) - c\,\partial_\theta B_{\rm w}^r]\frac{\boldsymbol{e}_\phi}{4\pi r}. \tag{17}$$

For short waves, it is helpful to view fields as functions of t, ξ, θ instead of t, r, θ. The fast oscillation then becomes isolated in the coordinate ξ. Differential equations can be rewritten in variables t, ξ, θ using

$$\partial_t|_r = \partial_t|_\xi + \partial_\xi|_t, \qquad \partial_r|_t = -\partial_\xi|_t. \tag{18}$$

Using the θ-component of induction Equation (14),

$$\partial_\xi(rE)_t = \partial_t(rB_{\rm w}^\theta)_\xi + \partial_\xi(rB_{\rm w}^\theta)_t, \tag{19}$$

we rewrite Equation (17) as

$$\boldsymbol{j} = [\partial_t(rE + rB_{\rm w}^{\theta})_{\xi} - c\,\partial_{\theta}B_{\rm w}^{r}]\frac{\boldsymbol{e}_{\phi}}{4\pi r}. \tag{20}$$

In short waves, the oscillation of $A_{\rm w}$ with ξ is much faster than its variation with t or θ at fixed ξ. Hence,

$$B_{\rm w}^{\theta} = \frac{1}{c\,r}\,\partial_{\xi}(rA_{\rm w}) \gg B_{\rm w}^{r}, \tag{21}$$

as $B_{\rm w}^{r}$ does not contain the large derivative $\partial_{\xi}(rA_{\rm w})$. Note also that $rE = \partial_{\xi}(rA_{\rm w})_t + \partial_t(rA_{\rm w})_{\xi}$, and

$$rE - rB_{\rm w}^{\theta} = \frac{1}{c}\,\partial_t(rA_{\rm w})_{\xi} \ll rE. \tag{22}$$

Therefore, the expression for $\boldsymbol{j}$ simplifies to

$$\boldsymbol{j} = \frac{\partial_t(rE)_{\xi}}{2\pi r}\,\boldsymbol{e}_{\phi}. \tag{23}$$

There is also a useful relation between plasma density $n = \rho/m$ and velocity $\boldsymbol{v}$ in short waves. When electron–positron pair creation is negligible inside the wave, n and $\boldsymbol{v}$ satisfy the continuity equation $\partial_t n + \nabla\cdot(n\boldsymbol{v}) = 0$. In the short-wave limit, it simplifies to

$$F_{\xi} \equiv (c - v_r)n = \text{const} = cn_{\rm bg}. \tag{24}$$

F_{ξ} is the particle flux through the surface of $\xi = \text{const}$. It is uniform across the wave packet, and evolves as the packet propagates to larger radii: $F_{\xi} = cn_{\rm bg} \propto r^{-3}$.

2.4. Wave–Plasma Interaction

We wish to find the evolution of the wave profile $E(\xi)$. It evolves because the electromagnetic field exchanges energy and momentum with the plasma oscillating in the wave. The plasma motion is described by four-velocity,

$$u^{\mu} = (\gamma, \boldsymbol{u}), \qquad \boldsymbol{u} \equiv \gamma\boldsymbol{\beta}, \qquad \gamma^2 = 1 + u^2 = \frac{1}{1-\beta^2}, \tag{25}$$

where $\boldsymbol{\beta} = \boldsymbol{v}/c$. The equation of motion is

$$\rho c^2\,\frac{d\boldsymbol{u}}{dt} = \boldsymbol{j}\times\boldsymbol{B}, \tag{26}$$

with the derivative d/dt taken along the fluid streamline: $d/dt = \partial_t + \boldsymbol{v}\cdot\nabla$. Taking the scalar product of both sides with $\boldsymbol{\beta}$, and using $\boldsymbol{E} + \boldsymbol{\beta}\times\boldsymbol{B} = 0$, one obtains

$$\rho c^2\,\frac{d\gamma}{dt} = \boldsymbol{E}\cdot\boldsymbol{j}. \tag{27}$$

This equation expresses energy conservation. Note that using $\rho c^2 = nmc^2$ in Equation (26) assumes a negligible contribution of internal (thermal) energy to the plasma inertial mass. As explained below, this is a reasonable approximation, because significant wave–plasma interaction happens where the plasma has a low temperature and a high γ.

Substitution of Equation (23) into Equation (27) gives

$$-\frac{E}{2\pi r^2}\,\partial_t(rE)_{\xi} = \rho c^2\,\frac{d\gamma}{dt} = \rho_{\rm bg}c^2\,\frac{d\gamma}{d\xi}, \tag{28}$$

where we used $d\xi = dt - dr/c = (1-\beta_r)dt$ along the fluid streamline and $\rho(1-\beta_r) = \rho_{\rm bg}$ (Equation (24)). The derivative $d\gamma/d\xi$ written out in coordinates $x^{\alpha} = (t, \xi, \theta, \phi)$ takes the form,

$$\frac{d\gamma}{d\xi} = \frac{dx^{\alpha}}{d\xi}\,\partial_{\alpha}\gamma = \frac{\partial_t\gamma}{1-\beta_r} + \partial_{\xi}\gamma + \frac{v_{\theta}\,\partial_{\theta}\gamma}{r(1-\beta_r)}. \tag{29}$$

One might expect the terms with $\partial_t\gamma$ and $\partial_{\theta}\gamma$ to be small compared to $\partial_{\xi}\gamma$ by the factor of $\sim\lambda/r \ll 1$. However, nonlinear evolution can make the term with $\partial_t\gamma$ important. As shown below, this happens when $\gamma^3 \gtrsim \sigma_{\rm bg}$.

Now we can write the energy equation in its final form:

$$\partial_t(r^2E^2) = -4\pi r^2\rho_{\rm bg}c^2\left[\partial_{\xi}\gamma + \frac{r\partial_t\gamma + \beta_{\theta}\partial_{\theta}\gamma}{r(1-\beta_r)}\right], \tag{30}$$

where derivative ∂_t is taken at fixed ξ, θ, ϕ, i.e., along the radial ray $r = ct + \text{const}$. This equation connects the evolution of $E(t, \xi, \theta)$ and $\gamma(t, \xi, \theta)$.

Equation (30) has a simple intuitive interpretation. Multiplication of both sides by an infinitesimal $\delta\xi$ gives

$$\partial_t\delta\mathcal{E}|_{\xi} = -\dot{\mathcal{N}}_{\xi}mc^2\delta\gamma. \tag{31}$$

Here, $\dot{\mathcal{N}}_{\xi} \equiv 4\pi r^2F_{\xi}$ is the isotropic equivalent of the particle flux F_{ξ}, and $\delta\gamma = (d\gamma/d\xi)\delta\xi$ is the change of γ along the fluid streamline as it crosses $\delta\xi$; it determines the energy gain of the plasma crossing $\delta\xi$ per unit time, $\dot{\mathcal{N}}_{\xi}mc^2\delta\gamma$. The quantity $\delta\mathcal{E} = c\,r^2E^2\,\delta\xi$ is the electromagnetic wave energy (isotropic equivalent) contained in the infinitesimal part $\delta\xi$ of the wave profile. This interpretation of $\delta\mathcal{E}$ is consistent with the energy density $U_{\rm w} = (E^2 + B_{{\rm w},\theta}^2 + B_{{\rm w},r}^2)/8\pi \approx E^2/4\pi$ for the electromagnetic wave in the short-wave limit.

Thus, Equation (30) merely states that the wave profile $r^2E^2(\xi)$ evolves by exchanging energy with the plasma. The plasma Lorentz factor $\gamma(\xi)$ oscillates, and $E(\xi)$ becomes deformed because the plasma passing through the wave receives energy at some ξ_1 and returns it to the electromagnetic field at another ξ_2. $E(\xi)$ is systematically reduced where $d\gamma/d\xi > 0$ and increased where $d\gamma/d\xi < 0$, leading to the steepening of $E(\xi)$.

The FFE limit of MHD would correspond to setting $\rho_{\rm bg} \to 0$, so that the right-hand side (RHS) of Equation (30) vanishes. Hence, the FFE solution is $rE = \text{const}$ at $\xi = \text{const}$. This wave excites no electric current $\boldsymbol{j}$, as can be verified using Equation (13), and so the FFE limit gives a simple vacuum wave superimposed on $\boldsymbol{B}_{\rm bg}$.

Equation (30) is sufficient to find the evolution of MHD waves if they drive pure radial plasma motions, as happens in the equatorial waves described below. Then, energy conservation (or the radial component of momentum Equation (26)) contains the complete dynamical information. Wave propagation outside the equatorial plane ($\theta \neq \pi/2$) involves additional θ-motions governed by the θ-component of Equation (26).

2.5. Equatorial Waves

The main features of shock development will be shown by tracking waves at $\theta = \pi/2$ that are symmetric about the equatorial plane. Symmetry implies $u^{\theta} = 0$ and $B^{r} = 0$ at $\theta = \pi/2$. The MHD fluid then oscillates with velocity

$$\boldsymbol{\beta} = \frac{\boldsymbol{E}\times\boldsymbol{B}}{\boldsymbol{B}^2} = \beta\,\boldsymbol{e}_r. \tag{32}$$

We will use the following notation:

$$\beta \equiv \beta_r, \qquad B \equiv B_{\theta}, \qquad E \equiv -E_{\phi}. \tag{33}$$

These definitions imply $E^2 = \boldsymbol{E}^2$, $B^2 = \boldsymbol{B}^2$ at $\theta = \pi/2$ while E, B, and β may be positive or negative. Note that

$$\beta = \frac{E}{B}, \qquad \gamma^2 = \frac{B^2}{B^2 - E^2}. \tag{34}$$

Equation (30) for the equatorial wave becomes

$$\partial_t(r^2E^2) = -4\pi r^2 \rho_{\rm bg} c^2 \left(\partial_\xi \gamma + \frac{\partial_t \gamma}{1-\beta} \right). \tag{35}$$

Recall that the partial derivatives ∂_t and ∂_ξ are defined in variables (t, ξ) and can act on any function $E(t, \xi)$, $\gamma(t, \xi)$, or $r(t, \xi) = c(t - \xi)$ entering Equation (35).

The equatorial wave problem can be reduced to one unknown function of t, ξ. Indeed, any short-wave packet propagating into an initially unperturbed plasma has a useful feature: fluid compression $n/n_{\rm bg}$ at any point in the packet is related to the local speed β (Equation (24)). The magnetic field is frozen in the fluid and compressed by the same factor, so

$$\frac{n}{n_{\rm bg}} = \frac{\rho}{\rho_{\rm bg}} = \frac{B}{B_{\rm bg}} = (1-\beta)^{-1}. \tag{36}$$

The electric field $E = \beta B$ can be expressed as

$$E = \frac{B_{\rm bg}\beta}{1-\beta}. \tag{37}$$

Thus, all MHD quantities in a short wave are functions of β, and Equation (35) can be recast so that it contains only derivatives of γ. This requires expressing $\partial_t(rE)$ in terms of $\partial_t\gamma$. Differentiating Equation (37), we find[5]

$$\partial_t(rE) = -2cE + \frac{rB_{\rm bg}\,\partial_t\gamma}{(1-\beta)^2\gamma^3\beta}, \tag{38}$$

where we used $\partial_t(rB_{\rm bg})_\xi = c\, d(rB_{\rm bg})/dr = -2cB_{\rm bg}$ and $d\gamma = \gamma^3\beta\, d\beta$. Then, substituting Equation (38) into Equation (35), we find

$$\left[\frac{2\sigma_{\rm bg}}{\gamma^3(1-\beta)^3} + \frac{1}{1-\beta} \right] \partial_t\gamma + \partial_\xi\gamma = \frac{4c\,\sigma_{\rm bg}\beta^2}{r(1-\beta)^2}. \tag{39}$$

Note that the term $(1-\beta)^{-1}$ in the bracket (brought by $\partial_t\gamma$ on the RHS of Equation (35)) is negligible only if $2\gamma^3 \ll \sigma_{\rm bg}$. This condition corresponds to $(\partial_t U_{\rm pl})_\xi \ll (\partial_t U_{\rm w})_\xi$ where $U_{\rm pl} = \gamma nmc^2$ and $U_{\rm w} = E^2/4\pi$.

A more convenient variable related to β is the compression of proper density $\tilde{\rho} = \rho/\gamma$,

$$\kappa \equiv \frac{\tilde{\rho}}{\rho_{\rm bg}} = \frac{\tilde{B}}{B_{\rm bg}} = \frac{\sigma}{\sigma_{\rm bg}} = \sqrt{\frac{1+\beta}{1-\beta}}, \tag{40}$$

where $\tilde{\rho}$ and $\tilde{B}$ are measured in the fluid rest frame, and $\sigma \equiv \tilde{B}^2/4\pi\tilde{\rho}c^2$. Note that $-1 < \beta < 1$ corresponds to $-\infty < \kappa < \infty$. Equation (39) rewritten in terms of κ becomes

$$\left(2\sigma_{\rm bg}\kappa^3 + \frac{1}{2} \right) \partial_t\kappa + \partial_\xi\kappa = \frac{2c}{r}\,\sigma_{\rm bg}\kappa^2(\kappa^2 - 1). \tag{41}$$

Here, in the coefficient of $\partial_t\kappa$ we neglected $\kappa^2/2 \ll 2\sigma_{\rm bg}\kappa^3$, using

$$2\sigma_{\rm bg}\kappa^3 + \frac{1+\kappa^2}{2} = 2\sigma_{\rm bg}\kappa^3 \left[1 + \mathcal{O}\left(\frac{1}{\sigma_{\rm bg}\kappa} \right) \right] + \frac{1}{2}. \tag{42}$$

The equatorial wave evolution will be found if we solve Equation (41) for $\kappa(t, \xi)$. This first-order partial differential equation is linear in the derivatives $\partial_t\kappa$ and $\partial_\xi\kappa$, and can be solved using the method of characteristics.

2.6. Characteristics

One can rewrite Equation (41) as

$$\left. \frac{d\kappa}{dt} \right|_{C^+} = \frac{c(\kappa - \kappa^{-1})}{r\,[1 + (4\sigma_{\rm bg}\kappa^3)^{-1}]}, \tag{43}$$

where the derivative $d\kappa/dt$ is taken along curves C^+ (characteristics) defined by

$$\frac{d\xi_+}{dt} = \frac{2}{4\sigma_{\rm bg}\kappa^3 + 1}. \tag{44}$$

The characteristics $\xi_+(t)$ can also be described by their radial speed,

$$\beta_+ = \frac{1}{c}\frac{dr_+}{dt} = 1 - \frac{d\xi_+}{dt} = \frac{4\sigma_{\rm bg}\kappa^3 - 1}{4\sigma_{\rm bg}\kappa^3 + 1}. \tag{45}$$

In an accompanying paper we give an alternative derivation of the wave evolution equation generalized to relativistically hot plasmas, and discuss two families of MHD characteristics $C^\pm$ propagating with radial speeds[6]

$$\beta_\pm = \frac{\beta \pm \beta_{\rm s}}{1 \pm \beta\beta_{\rm s}}. \tag{46}$$

Here, $\beta_{\rm s}$ is the wave speed relative to the plasma (the "magnetosonic speed"). In a magnetically dominated plasma, $\beta_{\rm s}$ is close to unity. In particular, for a cold plasma,

$$\beta_{\rm s}^2 = \frac{\sigma}{1+\sigma} = \frac{\kappa\sigma_{\rm bg}}{1+\kappa\sigma_{\rm bg}}. \tag{47}$$

The FFE limit corresponds to $\sigma_{\rm bg} \to \infty$ and $\beta_{\rm s} \to 1$. In this limit, Equations (43) and (44) simplify to $d\kappa/d\ln r = \kappa - \kappa^{-1}$ and $dr_+/dt = c$ ($\Leftrightarrow d\xi_+/dt = 0$), and give the solution

$$\kappa = \sqrt{1 + 2Kr^2}, \tag{48}$$

where $K = \text{const}$ along C^+. The corresponding solution for E is

$$E = \frac{\mu K}{r} \qquad \text{(FFE)}. \tag{49}$$

2.7. Shock Formation

Significant MHD corrections to FFE arise if $\sigma = \kappa\sigma_{\rm bg}$ drops, i.e., if the plasma experiences strong expansion in the wave, $\kappa \ll 1$. This occurs where the plasma is accelerated as E^2 approaches B^2. Note that strong expansion also implies strong adiabatic cooling. Thus, shock formation is expected in the part of

[5] The short-wave approximation for $E(\beta)$ (Equation (37)) is equivalent to $E = B_{\rm w}$. It can be used when expressing $\partial_t E|_\xi$ in terms of $\partial_t\beta|_\xi$. Note, however, that setting $E = B_{\rm w}$ would not be safe in expressions containing the large derivative ∂_ξ. In particular, $\partial_\xi(B_{\rm w} - E)$ in short waves is of the same order as $\partial_t E|_\xi$.

[6] C^- did not explicitly appear in our derivation above. For short waves, integration of MHD equations along C^- is just another way to get the relation between n, B, and β (Equation (36)), and the wave evolution problem is reduced to integration along C^+.

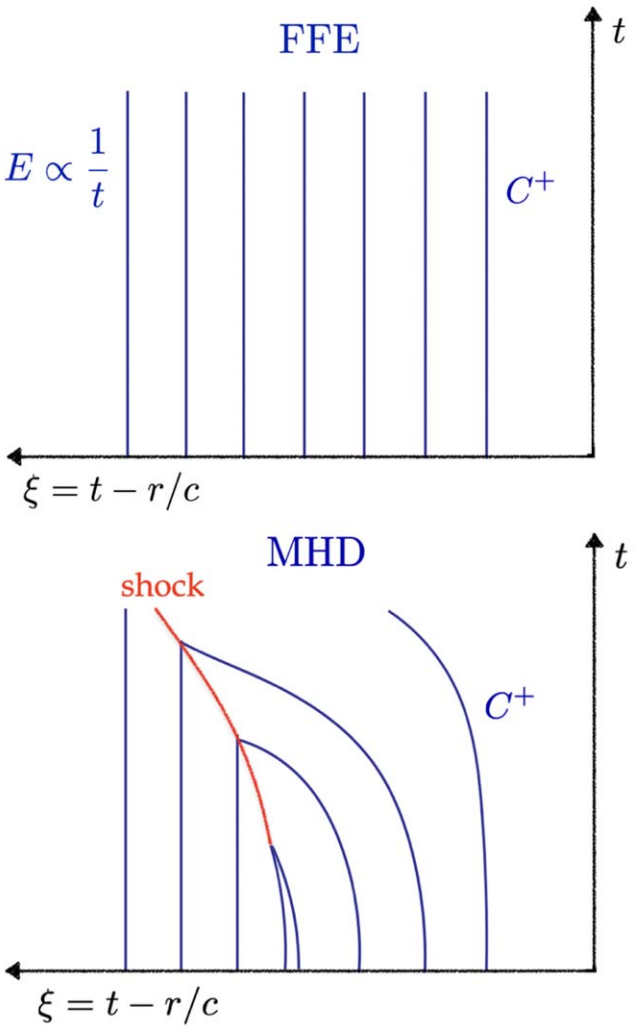

Figure 3. Schematic illustration of the flow of characteristics C^+ in FFE and MHD, shown in the t–ξ plane. Each C^+ has its initial position ξ_i and carries a value of $K(\xi_i) = rE/\mu$ determined by the initial wave profile. In the FFE limit, the characteristics are vertical straight lines; they propagate with speed $\beta_+ = 1$, which corresponds to $d\xi_+/dt = 0$. The MHD correction to FFE implies $d\xi_+/dt = 1 - \beta_+ \neq 0$, so the C^+ characteristics are no longer static in ξ; they become significantly bent in the region where E^2 approaches B^2, leading to the formation of a shock (red curve).

the wave with large γ, small κ, and a reduced temperature. The process of MHD shock formation is illustrated in Figure 3. The shock occurs where the characteristics cross, bringing different values of β to the same location and thus creating a discontinuity.

We consider waves emitted at sufficiently small radii where B_{bg} far exceeds the wave electric field E, and the plasma oscillates with small $|\beta| \ll 1$, which implies a negligible change in plasma density, $|\kappa - 1| \ll 1$. At the small radii, characteristics propagate with speed $\beta_+ = 1 - \mathcal{O}(\sigma_{bg}^{-1})$, and FFE is an excellent approximation. In particular, the wave initially propagates with negligible distortion: each C^+ characteristic satisfies $d\xi_+/dt = 0$, i.e., keeps a constant coordinate $\xi = t - r/c = \xi_i$, and one can define an initial (undeformed) profile of $E(\xi_i)$. More precisely, the profile of $rE(\xi)$ is static at small r while the normalization of $E(\xi)$ is decreasing, $E \propto r^{-1}$.

This initial profile is the only parameter of the problem other than the magnetospheric parameters $\mu = r^3 B_{bg}$ and $D = r^3 \sigma_{bg}$. It is conveniently described by

$$K(\xi_i) \equiv \frac{rE}{\mu} \quad \text{(set at small } r\text{)}. \tag{50}$$

We will calculate the wave evolution by tracking the C^+ characteristics, each described by its initial position ξ_i and $K(\xi_i)$. Note that $K(\xi_i)$ may be positive or negative. Characteristics with $K < 0$ will develop $\kappa \ll 1$, leading to monster shock formation.

It is easy to see that even an arbitrarily high $\sigma_{bg} \to \infty$ does not save the MHD wave from breaking. Indeed, consider a wave with E oscillating between $\pm E_0$. As the wave expands from small radii, where $E_0(r) \ll B_{bg}(r)$, the ratio $E_0/B_{bg} \propto r^2$ grows and eventually approaches $1/2$ at some radius $R_\times$. At this moment, the minimum $E = -E_0 = -B_{bg}/2$ and

$$E = -B, \quad \beta = -1, \quad \gamma \to \infty, \quad \kappa \to 0. \tag{51}$$

An arbitrarily high σ_{bg} does not prevent the decrease of $\sigma = \kappa\sigma_{bg}$, which reduces the speed of C^+ characteristics, $dr_+/dt < c$. The characteristics become bent and eventually collide, forming a shock (Figure 3).

A formal proof of shock formation is provided by the solution of the coupled Equations (43) and (44) for $\kappa(t)$ and $\xi_+(t)$ or $r_+(t) = c(t - \xi_+(t))$ (Appendix A). Numerical examples will be given in Section 4.

3. Method for Tracking Shocks

3.1. Shock Strength

The plasma speed β is discontinuous at the MHD shock because the upstream and downstream characteristics bring to the shock different values of β: $\beta_u \neq \beta_d$ (subscripts "u" and "d" refer to the immediate upstream and downstream of the shock). The corresponding jump in the proper density characterizes the shock strength,

$$q \equiv \frac{\tilde{\rho}_d}{\tilde{\rho}_d} = \frac{\kappa_d}{\kappa_u} = \frac{\gamma_d(1 + \beta_d)}{\gamma_u(1 + \beta_u)}. \tag{52}$$

It is related to the Lorentz factor of the upstream plasma relative to the downstream, $\Gamma_{u|d} = \gamma_u\gamma_d(1 - \beta_u\beta_d)$. A similar expression for $\Gamma_{u|d}$ holds in the shock rest frame $\mathcal{K}'$. To distinguish between different frames, dynamical quantities measured in the shock frame will be denoted with a prime, and quantities measured in the drift frame $\tilde{\mathcal{K}}$ (in which $\tilde{E} = 0$) are denoted with a tilde. Using $\gamma'_u, \gamma'_d \gg 1$, one can write the continuity of mass flux as $\tilde{\rho}_u\gamma'_u \approx \tilde{\rho}_d\gamma'_d$, and find

$$\Gamma_{u|d} \approx \frac{1}{2}\left(\frac{\gamma'_u}{\gamma'_d} + \frac{\gamma'_d}{\gamma'_u}\right) \approx \frac{1}{2}\left(q + \frac{1}{q}\right). \tag{53}$$

3.2. Shock Speed

We can track the evolution of waves with shocks if we know the shock Lorentz factor relative to the downstream plasma $\gamma_{sh|d}$. The standard jump conditions for perpendicular relativistic magnetized shocks give $\gamma_{sh|d} \approx \sqrt{\sigma_u} \gg 1$. The shock is collisionless and mediated by Larmor rotation, so the jump normally occurs on the Larmor scale, as verified by detailed kinetic simulations (Sironi et al. 2021). Dissipation of the upstream kinetic energy occurs through gyration and

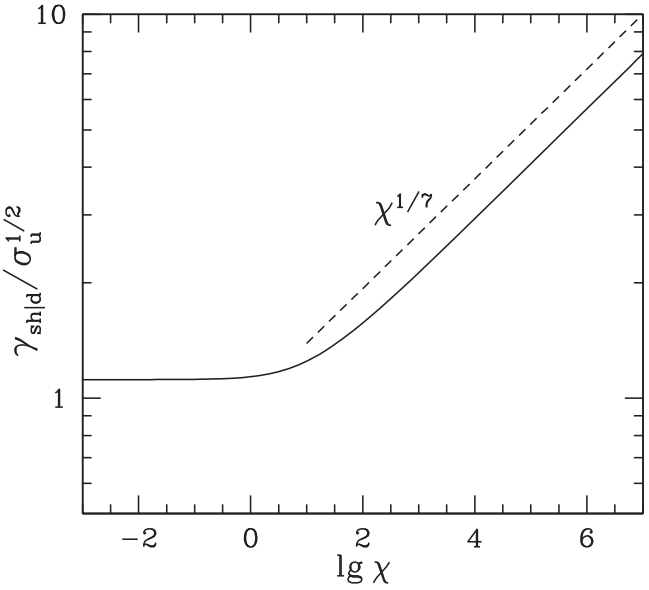

Figure 4. Lorentz factor of the shock relative to the downstream plasma as a function of the radiative parameter χ (solid curve, calculated as explained in Appendix B). The dashed line shows the analytical result at $\chi \gg 1$ (Equation (54)).

thermalization enabled by the instability of coherent gyration in the downstream.

However, the standard picture of Larmor-mediated shocks can fail for the monster shocks, where particles experience extremely fast radiative losses. The ultrarelativistic plasma flow entering the shock is subject to radiation reaction in response to curvature in the particle trajectory. The upstream flow may experience a strong radiative drag before it completes one Larmor rotation and joins the downstream.

The resulting jump conditions can be evaluated by going beyond the MHD description and examining microphysics of the shock transition on scales smaller than the particle Larmor radius. This is done in Appendix B, where we calculate the structure of the flow across the jump numerically and also derive an approximate solution analytically. The result is shown in Figure 4, and the jump condition may be stated as

$$\gamma_{\rm sh|d} \approx (1+\chi)^{1/7}\sqrt{\sigma_{\rm u}}, \tag{54}$$

which smoothly matches the results at $\chi \ll 1$ and $\chi \gg 1$. The parameter χ is defined in Equation (B9) in terms of the upstream Lorentz factor $\gamma'_{\rm u}$ and magnetic field $B'_{\rm u}$ measured in the shock frame. Using $B'_{\rm u} = \gamma'_{\rm u}\tilde{B}_{\rm u}$ and $\gamma'_{\rm u} = q\gamma'_{\rm d} = q\gamma_{\rm sh|d}$, one can express it as

$$\chi = \frac{\sigma_{\rm T}\tilde{B}_{\rm u}q^3\gamma_{\rm sh|d}^3}{4\pi e\,\sigma_{\rm u}^{3/2}}, \tag{55}$$

where $\tilde{B}_{\rm u} = \kappa_{\rm u}B_{\rm bg}$ and $\sigma_{\rm u} = \kappa_{\rm u}\sigma_{\rm bg}$. Equations (54) and (55) can be solved for $\gamma_{\rm sh|d}$ and χ, which gives

$$\chi^{4/7} \approx \frac{\sigma_{\rm T}B_{\rm bg}\kappa_{\rm d}^3}{4\pi e\kappa_{\rm u}^2}. \tag{56}$$

This expression is found for $\chi \gg 1$, and the solution for $\chi < 1$ is not needed, as its value makes no difference for $\gamma_{\rm sh|d} \approx (1+\chi)^{1/7}\sqrt{\sigma_{\rm u}} \approx \sqrt{\sigma_{\rm u}}$.

The shock Lorentz factor in the lab frame is

$$\gamma_{\rm sh} = \gamma_{\rm sh|d}\gamma_{\rm d}(1+\beta_{\rm sh|d}\beta_{\rm d}) \approx \gamma_{\rm sh|d}\kappa_{\rm d}, \tag{57}$$

where we used $1+\beta_{\rm sh|d}\beta_{\rm d} \approx 1+\beta_{\rm d}$, because $\gamma_{\rm sh|d} \gg \gamma_{\rm d}$. The shock motion in the ξ-coordinate is described by $d\xi_{\rm sh}/dt = 1-\beta_{\rm sh} \approx (2\gamma_{\rm sh}^2)^{-1}$. Thus, we find

$$\frac{d\xi_{\rm sh}}{dt} \approx \frac{1}{2\gamma_{\rm sh}^2} \approx \begin{cases} \left(\dfrac{\pi e}{\sigma_{\rm T}B_{\rm bg}\sigma_{\rm bg}^2\kappa_{\rm d}^7}\right)^{1/2} & \chi > 1 \\ (2\sigma_{\rm bg}\kappa_{\rm u}\kappa_{\rm d}^2)^{-1} & \chi < 1. \end{cases} \tag{58}$$

There is one caveat in the derivation of this result: we neglected that the shock can generate an electromagnetic precursor, which interacts with the upstream flow and may reduce its Lorentz factor $\gamma_{\rm u}$ (Section 5.5). Note however that $d\xi_{\rm sh}/dt$ is independent of $\kappa_{\rm u} \approx (2\gamma_{\rm u})^{-1}$ when $\chi \gg 1$. Furthermore, the shock speed will be dynamically regulated to a value independent of $\kappa_{\rm u}$ and $\kappa_{\rm d}$ when $\kappa_{\rm d} \ll 1$, as explained in Section 5.4.

3.3. Tracking Waves with Shocks

The MHD wave evolution is controlled by the flow of C^+ characteristics in coordinate $\xi = t - r/c$. This flow is described by Equations (43) and (44). They determine both the shape of characteristics and the values of κ, β, and E on each C^+.

After caustic formation, the born shock separates the C^+ flow into two regions: upstream and downstream (ahead and behind the shock). The characteristics propagate with different speeds in these regions and collide at the shock. The colliding characteristics $C_{\rm u}^+$ and $C_{\rm d}^+$ are terminated and disappear from the wave evolution problem. The location of $C_{\rm u}^+$-$C_{\rm d}^+$ collision moves with speed $\beta_{\rm sh}$ determined by Equation (58). The shock always propagates faster than $C_{\rm u}^+$ and slower than $C_{\rm d}^+$: $\beta_{\rm u}^+ < \beta_{\rm sh} < \beta_{\rm d}^+$.

Equations (43), (44), and (58) give a closed description for MHD waves with shocks. The cold approximation $\gamma_{\rm s}^2 = \sigma$ used in Equations (44) turns out to be sufficient for the kilohertz waves studied in this paper. The accurate $\gamma_{\rm s} < \infty$ is important in the pre-shock region where κ drops so much that the C^+ flow experiences significant deformation $d\xi_+/dt$ (Figure 3). This region is also coldest (due to adiabatic cooling accompanying the drop in κ), and so $\gamma_{\rm s}^2 = \sigma$. In the post-shock region $\xi > \xi_{\rm sh}$, fast radiative losses also allow one to use $\gamma_{\rm s}^2 \approx \sigma$. Here, the evolution is simple anyways, not sensitive to the precise $\gamma_{\rm s}$: the post-shock $\gamma_{\rm s}$ is so high that C^+ remain practically static in the ξ coordinate during the main shock dissipation phase $r \lesssim 3R_\times$.

3.4. Numerical Implementation

The described method for calculating the wave evolution is easily implemented in a numerical simulation. When launching a wave, we set up an initially uniform grid in $\xi_{\rm i}$ of size N_+, and then use the N_+ characteristics to track the wave evolution in the ξ coordinate. In the simulations presented below, we used $N_+ = 10^5$. At each time step dt, the displacement $d\xi_+$ of each characteristic and the change of its κ are determined by Equations (44) and (43). After each time step, the code examines the updated positions or the characteristics and checks for their crossing to detect shock formation. Once the shock is born, the code begins to track its motion according to

Equation (58) and also check at each time step which characteristics terminate at the shock.

The simulation thus follows the entire evolution of the wave, from its initial deformation-free propagation at $r < R_\times$ to shock formation at $r \approx R_\times$ to subsequent evolution with the embedded shock. We use an adaptive time step to resolve fast changes in MHD quantities that occur near $R_\times$. Note also that the density of characteristics drops ahead of the shock, where κ is lowest and $d\xi_+/dt$ is highest. To maintain sufficient spatial resolution, we use adaptive mesh refinement in ξ_i without changing the total number N_+ of active (not terminated) characteristics. This is achieved by launching new characteristics in the region of low resolution while discarding the characteristics terminated at the shock.

The calculated $\xi_+(t)$ and $\kappa(t)$ along each C^+ determine the wave profile $\kappa(\xi)$ at any time t. Thus, we find the evolution of the wave profile.

4. Sample Numerical Models

As a concrete example, consider a wave with the initial profile $E(\xi_i) = E_0 \sin(\omega\xi_i)$. Recall that $E \propto r^{-1}$ along each characteristic C^+ until it develops an extremely low compression factor κ where $B^2 - E^2$ approaches zero. The initial profile of E sets the parameter $K = rE/\mu$ on each C^+,

$$K(\xi_i) = K_0 \sin(\omega\xi_i), \qquad K_0 = \frac{rE_0(r)}{\mu}. \tag{59}$$

The wave trough $E = -E_0$ approaches $-B_{bg}/2$ (so E^2 approaches B^2) at radius $R_\times$. It is related to K_0 by

$$R_\times \approx \frac{1}{\sqrt{2K_0}}. \tag{60}$$

Radius $R_\times$ is also related to the isotropic equivalent of the wave power $L \approx cr^2E_0^2/2$ and the magnetic dipole moment of the star μ:

$$R_\times = \left(\frac{c\mu^2}{8L}\right)^{1/4} \approx 1.4 \times 10^8 \frac{\mu_{33}^{1/2}}{L_{43}^{1/4}} \text{ cm}. \tag{61}$$

The background magnetization parameter at $R_\times$ is determined by Equation (8),

$$\sigma_\times \equiv \sigma_{bg}(R_\times) \approx 3.6 \times 10^9 \frac{\mu_{33}^{1/2} L_{43}^{3/4}}{\mathcal{N}_{37}}. \tag{62}$$

Our sample models assume the plasma density parameter $\mathcal{N} = 10^{39}$ and 10^{37} (Section 2.1). In both models, the magnetosphere has dipole moment $\mu = 10^{33}\,\text{G cm}^3$. The wave has frequency $\nu = \omega/2\pi = 10$ kHz and initial power $L = 10^{43}$ erg s^{-1}. The simulation results are shown in Figure 5 and may be summarized as follows.

The wave travels with no distortion until it comes to $R_\times$ where E^2 nearly reaches B^2. The plasma at this point develops an ultrarelativistic drift with $u \approx -(1/2)\sqrt{B_{bg}/(B - |E|)}$, leading to immediate shock formation. In particular, in the model with $\mathcal{N} = 10^{39}$, the plasma four-velocity upstream of the shock reaches $u_u \sim -10^5$, and the shock compression factor $q = \kappa_d/\kappa_u$ exceeds 10^4. The shock has a large radiative parameter $\chi \gg 1$ (Figure 6). An analytical derivation reproducing the numerical results will be given in Section 5. We will show there that the Lorentz factor of the accelerated plasma scales as $\gamma \propto L/\mathcal{N}\nu$, and our second sample model (with $\mathcal{N} \sim 10^{37}$) gives so high γ that the MHD description breaks and the plasma dynamics needs a two-fluid description.

As one can see in Figure 5, the wave profile $E(\xi)$ develops a plateau of $E \approx -B_{bg}/2$, which corresponds to $E^2 \approx B^2$. The small variation δE along the plateau implies that the plateau is formed by stretching a small interval $\delta\xi$ of the initial profile $E(\xi)$. The stretching is clearly demonstrated by the C^+ flow. The part of the wave profile where E^2 approaches B^2 develops a low $\kappa^2 = (B - |E|)/B_{bg}$, and here $d\xi_+/dt$ steeply increases (Equation (44)). As a result, characteristics with $E^2 \approx B^2$ swiftly "fall" toward the shock, with acceleration, like a waterfall. This effect protects the MHD wave from breaking the condition $E^2 < B^2$. Where κ becomes so low that $\kappa^3\sigma_{bg} \ll 1$, the C^+ motion in ξ saturates at $d\xi_+/dt \approx 2$. The fast, large displacement of C^+ in ξ creates a nearly perfect plateau of $E \approx -B \approx -B_{bg}/2$, with a small difference $B - |E| \approx \kappa^2 B_{bg}$.

As the wave propagates to larger radii and $B_{bg} \propto r^{-3}$ decreases, the plateau level $E = -B_{bg}/2$ moves up and away from the original minimum of the sine wave ($E = -E_0$), and hence its width grows, so the plateau is forced to occupy an increasing part of the wave profile. At radii $r \gg R_\times$, where $B_{bg}/2E_0 = R_\times^2/r^2 \ll 1$, the plateau approaches $E \approx 0$, i.e., the part of the wave with $E < 0$ becomes erased. As a result, the wave loses about half of its original energy.[7] Waves with multiple oscillations lose even more than half: one can see that the shock formed in the first oscillation at $\xi \approx 3\pi/2\omega$ eventually enters the second oscillation $\xi > 2\pi/\omega$ before stalling there, so the final plateau of $E = 0$ occupies slightly more than half of the oscillation.

The plasma speed $\beta = E/B = E/(B_{bg} + E)$ is positive where $E > 0$ and reaches maximum at the wave crest. This maximum β approaches unity at $r \gg R_\times$. Here, the plasma speed relative to the wave, $1 - \beta \approx B_{bg}/E$, becomes small, increasing the time it takes the plasma to cross the first half of the wave oscillation,

$$t_{cross} = \int_0^{\pi/\omega} \frac{d\xi}{1 - \beta} \approx \frac{r^2}{\omega R_\times^2} \quad (r \gg R_\times). \tag{63}$$

The short-wave approximation $t_{cross} \ll r/c$ holds throughout our simulation. The simulation did not follow how the plasma eventually becomes trapped in the wave, $t_{cross} > r/c$; this occurs later, when the wave propagates to larger radii.

Besides the monster shock launched at $\xi \approx 3\pi/2\omega$ at radius $R_\times$, the simulation shows gradual steepening of the wave at the leading edge $\xi = 0$ (and at $\xi \approx 2\pi/\omega$). At radius $R_F \approx 10R_\times$, the wave launches a forward shock as shown in Section 5.6.

5. Analytical Description

The wave evolution demonstrated by the numerical simulation may also be derived analytically. Two key dimensionless parameters of the problem are

$$\zeta \equiv \frac{\omega R_\times}{c\,\sigma_\times} = \pi^2 \frac{mc^2\mathcal{N}\nu}{L} \approx 8 \times 10^{-8}\,\mathcal{N}_{37}L_{43}^{-1}\nu_4, \tag{64}$$

$$\eta \equiv \frac{\sigma_T B_\times}{8\pi e} = \frac{r_e\omega_\times}{3c} \approx 2 \times 10^{-8}\,\mu_{33}^{-1/2}L_{43}^{3/4}, \tag{65}$$

where $B_\times \equiv B_{bg}(R_\times)$ and $\omega_\times \equiv eB_\times/mc$.

[7] Maximum dissipation fraction of the shock is $\sigma_u^{-1} \sim c/\omega R_\times \ll 1$. This small factor is compensated by the large ratio of the traveled distance $\sim R_\times$ to c/ω, resulting in strong dissipation.

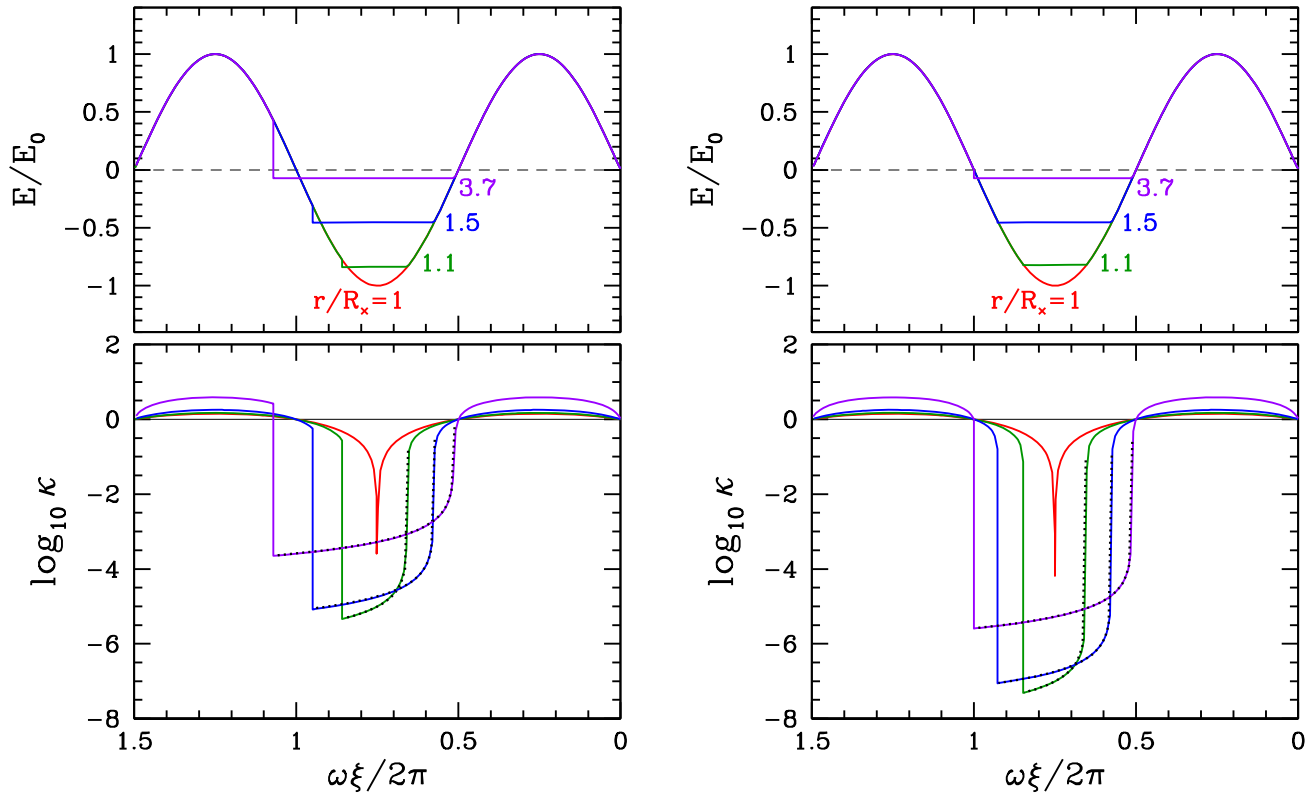

Figure 5. Left: evolution of the wave profile $E(\xi)$ and $\kappa(\xi)$, where $\xi \equiv t - r/c$. The wave has frequency $\nu = \omega/2\pi = 10$ kHz and initial power $L = 10^{43}$ erg s^{-1}; the magnetosphere has magnetic dipole moment $\mu = 10^{33}$ G cm^3 and density parameter $\mathcal{N} = 10^{39}$. Five snapshots are shown, when the wave packet reaches $r/R_\times = 0.9$ (black), 1 (red), 1.1 (green), 1.5 (blue), and 3.7 (magenta). The electric field E is normalized to E_0, which would be the wave amplitude if it propagated in vacuum. The plasma Lorentz factor γ is related to the proper compression κ by $\gamma = (1 + \kappa^2)/2\kappa$. Black dotted curves show the analytical result for $\kappa(\xi)$ (Equation (78)). The simulation neglected the shock precursor effect, which can reduce γ in the interval $3\pi/2\omega < \xi < \xi_{sh}$ (Section 5.5). Right: same wave but now launched into the magnetosphere with $\mathcal{N} = 10^{37}$. Here, κ becomes extremely small, breaking the MHD description and transitioning to a two-fluid regime. We argue in Section 5.3 that the two-fluid calculation will likely give the same evolution of $E(\xi)$ and $\kappa(\xi)$ as found in the MHD model.

5.1. Shock Formation: Caustic in the C^+ Flow

Each C^+ starts at small radii with $\xi_+ = \xi_i$, and one can think of ξ_i as a Lagrangian coordinate in the flow of characteristics $\xi_i \to \xi_+(\xi_i, t)$ with initial condition $\xi_+(\xi_i, 0) = \xi_i$. The C^+ flow satisfies Equation (A11) found from Equations (43) and (44):

$$\xi_+ = \xi_i + \frac{(\kappa - 1)^2}{8cDK^2\kappa} = \xi_i + \frac{\gamma - 1}{4cDK^2}, \tag{66}$$

where $D = \sigma_{bg} r^3 = \sigma_\times R_\times^3$. The caustic in the C^+ flow (birth of the shock) appears where $(\partial\xi_+/\partial\xi_i)_t$ vanishes. This means that characteristics carrying different values of K begin to cross, creating a discontinuity. This happens at some $\xi_i = \xi_i^c$ and time $t = t_c$, which may be derived analytically (Appendix A). We also find the plasma compression factor at the caustic, κ_c.

In the leading order of the small parameters ζ and $(R_\times\omega/c)^{-1}$, we obtain

$$\kappa_c = \frac{\sqrt{\zeta}}{24^{1/4}}, \qquad \cos(\omega\xi_i^c) = \frac{4\sqrt{\zeta}}{24^{3/4}} - \frac{c}{R_\times\omega}, \tag{67}$$

$$\frac{ct_c}{R_\times} - 1 = -\frac{\zeta}{3\sqrt{24}} - \frac{8\sqrt{\zeta}}{24^{3/4}}\frac{c}{R_\times\omega} + \frac{c^2}{4R_\times^2\omega^2}. \tag{68}$$

Powerful kilohertz waves have $\sqrt{\zeta} \ll c/R_\times\omega$. In this regime, C^+ characteristics turn back toward the star ($\beta_+ < 0$) before forming the caustic. The caustic occurs at ξ_c near $3\pi/2\omega$:

$$\omega\xi_c \approx \frac{3\pi}{2} + \frac{16\sqrt{\zeta}}{24^{3/4}} - \frac{c}{R_\times\omega}. \tag{69}$$

5.2. Plasma Motion Upstream of the Shock

As one can see in Figure 5, the plasma develops a very low compression factor κ ahead of the shock, a result of huge plasma acceleration toward the star. The acceleration occurs along the plateau of $E \approx -B_{bg}/2$, and the four-velocity u of the plasma drift in the wave is controlled by how close $B^2 - E^2$ approaches zero:

$$u \approx -\gamma \approx -\frac{B}{\sqrt{B^2 - E^2}} \approx -\frac{1}{2}\sqrt{\frac{B_{bg}}{B_{bg} + 2E}}. \tag{70}$$

The plasma acceleration is accompanied by its expansion by the factor of $\kappa^{-1} \approx 2\gamma$, as follows from the relation

$$\frac{\kappa^2 + 1}{2\kappa} = \gamma, \qquad \frac{\kappa^2 - 1}{2\kappa} = u. \tag{71}$$

It is the drop of κ that boosts the motion of C^+ characteristics in the ξ-coordinate, deforming the wave profile. The relation between κ and the displacement of characteristics $\xi_+ - \xi_i$ is given by Equation (66). Substituting $K = K_0 \sin(\omega\xi_i)$ and

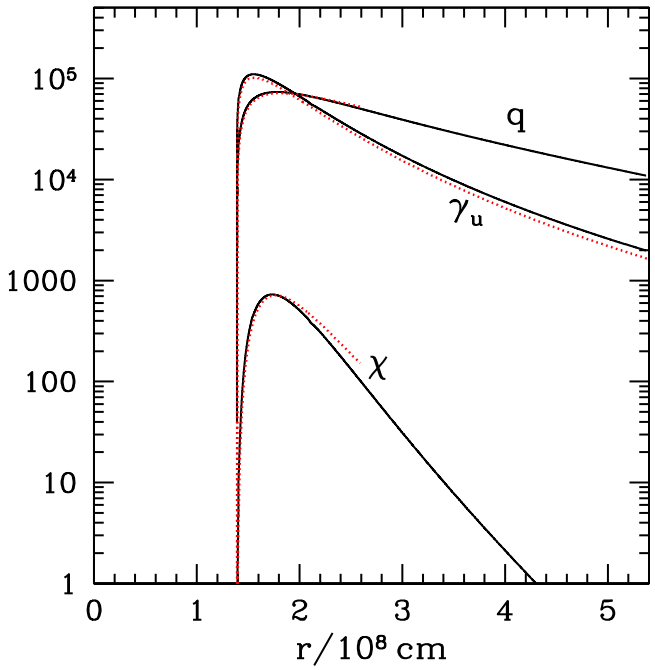

Figure 6. Evolution of shock parameters in the simulation with $\mathcal{N} = 10^{39}$: upstream Lorentz factor γ_u, shock compression factor $q \equiv \kappa_d/\kappa_u$, and radiative parameter χ. Red dotted curves show the analytical results for $\gamma_u(r)$, $q(r)$, $\chi(r)$ derived in Section 5 (Equations (75), (92), and (94)). The simulation neglected the precursor effect (Section 5.5).

$K_0^{-1} = 2R_\times^2$, we find

$$\xi_+ - \xi_i = \frac{(1-\kappa)^2 R_\times}{2c\,\kappa\sigma_\times \sin^2(\omega\xi_i)}. \tag{72}$$

One can use this relation to evaluate κ_u on the characteristic C_u^+ that reaches the shock at time t: $\xi_+^u = \xi_{sh}$. We are interested in waves that develop $\kappa_u \ll 1$, and then Equation (72) gives

$$\kappa_u \approx \frac{\zeta}{2\,\omega\Delta\xi\,\sin^2(\omega\xi_i^u)}, \tag{73}$$

where $\Delta\xi \equiv \xi_+^u - \xi_i^u$ is the displacement of C_u^+.

For the nascent shock at the caustic, we find $\omega\Delta\xi \approx 12\sqrt{\zeta}/24^{3/4}$, which corresponds to $\kappa_u = \kappa_c$ (Section 5.1). Later, $\Delta\xi(t)$ approximately equals the width of the plateau,

$$E \approx E_p = -\frac{B_{bg}}{2} \qquad \text{(plateau)}, \tag{74}$$

along which κ drops to κ_u (Figure 5). All C^+ populating the plateau have approximately the same E and so nearly the same ξ_i—the plateau forms by the huge stretching (by the factor of $\sim\kappa^{-1}$) of an initially small interval in the original wave profile $\delta\xi_i$. This stretching occurs because the characteristics "fall" onto the shock from ξ_i where E approaches $-B_{bg}/2$, i.e., $\sin(\omega\xi_i) \approx -B_{bg}/2E_0$.

The plateau begins and ends where $\sin(\omega\xi) \approx -B_{bg}/2E_0 \approx -R_\times^2/r^2$, so its width is $\omega\Delta\xi \approx \pi - 2\arcsin(R_\times^2/r^2)$. Then, from Equation (73) we find the plasma Lorentz factor just upstream of the shock,

$$\gamma_u \approx \frac{1}{2\kappa_u} \approx \frac{1}{\zeta x^4}\left(\pi - 2\arcsin\frac{1}{x^2}\right), \quad x \equiv \frac{r}{R_\times}. \tag{75}$$

This expression accurately reproduces the simulation results (Figure 6). The maximum $\gamma_u \approx \zeta^{-1}$ is reached at $r/R_\times \approx (1+1/\pi)^{1/2} \approx 1.15$.

The same γ_u can be found from energy conservation, as the flow is accelerated at the expense of the electromagnetic wave energy $\mathcal{E} = \int c\,r^2E^2\,d\xi$ (isotropic equivalent). The profile $E(\xi)$ includes the plateau part ($E \approx E_p = -B_{bg}/2$) and the part where $E(\xi)$ is practically unchanged from the vacuum propagation $E(\xi) \approx E_0\sin(\omega\xi)$ (Figure 5). The plateau has the width $\Delta\xi$ and ends at the shock ξ_{sh} with a jump $[E^2] = E_d^2 - E_u^2$, where $E_u \approx E_p$. During time $dt \approx dr/c$, the "area" associated with the integral $\int r^2E^2d\xi$ changes: (1) the rise of the plateau erases a horizontal stripe $|d(r^2E_p^2)|\,\Delta\xi$ and (2) the motion of the shock erases a vertical stripe $r^2[E^2]\,d\xi_{sh}$. So, the wave loses energy

$$d\mathcal{E}_{lost} \approx -c\,d(r^2E_p^2)\Delta\xi + cr^2[E^2]\,d\xi_{sh}. \tag{76}$$

The energy lost at $\xi < \xi_{sh}$ (the first term on the RHS) goes into plasma acceleration upstream of the shock,

$$\gamma_u mc^2 d\mathcal{N} \approx 4c\,r^2E_p^2\,d\ln r\,\Delta\xi, \tag{77}$$

where $d\mathcal{N} = 4\pi r^2 cn_{bg}\,dt = 4\pi c\mathcal{N}d\ln r$ is the particle number (isotropic equivalent) passing through the wave during dt, and we used the fact that $E_p^2r^2 \propto r^{-4}$. Equations (77) and (75) give the same γ_u.

The consideration of the C^+ flow or energy conservation gives a simple solution for $\gamma(\xi)$ along the plateau. In particular, using $\kappa \ll 1$ and $\sin^2(\omega\xi_i) \approx R_\times^4/r^4$, we find from Equation (72),

$$\gamma(\xi) \approx \frac{1}{2\kappa} \approx \frac{c\sigma_{bg}}{r}(\xi - \xi_i) \qquad (\xi_i < \xi < \xi_{sh}). \tag{78}$$

The solution reproduces $\gamma(\xi)$ found in the simulations (Figure 5). It ends at the shock with $\gamma(\xi_{sh}) = \gamma_u$.

Note that the plasma magnetization in the wave $\sigma = \kappa\sigma_{bg}$ is lowest just upstream of the shock, $\sigma_u = \kappa_u\sigma_{bg} \approx \sigma_{bg}/2\gamma_u$. Its minimum value is reached soon after the wave crosses $R_\times$: $\sigma_{min} \sim \zeta\sigma_{bg} = \omega R_\times/c \gg 1$. This minimum value is independent of σ_{bg}.

5.3. Beyond Single-fluid MHD

The accelerated flow experiences dramatic expansion in the region $\xi_{sh} < \xi < \xi_i$, by a factor of $\kappa^{-1} \gg 1$, and the gyrofrequency in the fluid frame $\tilde{\omega}_B = \kappa\omega_B$ drops. Note that the proper time $\tilde{t}$ (measured in fluid frame) slows down by a similar factor, $d\tilde{t} = dt/\gamma \approx 2\kappa dt$.

Consider now the profile of $\gamma(\xi) \gg 1$ (Equation (78)). A fluid element crosses $d\xi$ in time $dt = d\xi/(1-\beta) \approx d\xi/2$, and we find that γ changes along the fluid streamline on the proper timescale $\tilde{t}_{ev}$ given by

$$\tilde{t}_{ev}^{-1} \equiv \frac{d\ln\gamma}{d\tilde{t}} = \frac{d\gamma}{dt} \approx 2\frac{d\gamma}{d\xi} \approx \frac{2c\sigma_{bg}}{r}. \tag{79}$$

MHD description assumes that the gyration timescale $\tilde{\omega}_B^{-1} \approx (\omega_B/2\gamma)^{-1}$ is much shorter than $\tilde{t}_{ev}$. This condition may be stated as

$$\tilde{\omega}_B\tilde{t}_{ev} \approx \frac{\omega_B r}{4\gamma\sigma_{bg}c} = \frac{\omega_\times r}{4\gamma\sigma_\times c} \gg 1. \tag{80}$$

It requires $\gamma \ll \gamma_{\rm MHD}$, where

$$\gamma_{\rm MHD} \approx \frac{\omega_\times r}{4\sigma_\times c} = \frac{\pi e \mathcal{N} r}{\mu}. \tag{81}$$

The MHD solution for the wave propagation problem remains mathematically well behaved even when γ reaches arbitrary high values. However, its applicability to a real plasma is vindicated only if $\gamma \ll \gamma_{\rm MHD}$, which roughly corresponds to $\zeta \gg (2\omega/\omega_\times)^{1/2}$. Our sample numerical model with $\mathcal{N} = 10^{39}$ had the peak $\gamma \approx 10^5$, just below $\gamma_{\rm MHD} \approx 2 \times 10^5$, so the MHD description is marginally valid near the peak and accurate at other radii. The model with $\mathcal{N} = 10^{37}$ developed $\gamma \gg \gamma_{\rm MHD}$, breaking the MHD condition.

In waves with $\zeta \lesssim (2\omega/\omega_\times)^{1/2}$, where γ reaches $\gamma_{\rm MHD}$, the e^+ and e^- motions are no longer the common $\boldsymbol{E} \times \boldsymbol{B}$ drift, and they are no longer coupled to the magnetic field via gyration. The momenta of unmagnetized e^+ and e^- become significantly different and a two-fluid description is required to formulate the electric current $\boldsymbol{j}$ and its effect on the electromagnetic field of the wave. Main features of the two-fluid solution may be anticipated using the following two considerations.

(1) The symmetry of e^+ and e^- motions in an electromagnetic wave implies that they have equal velocities β_z along the wavevector $\boldsymbol{k}$ and opposite transverse velocities $\pm\beta_x$ with $\beta_y = 0$ (Beloborodov 2022). We here use local Cartesian coordinates x, y, z with the z-axis along $\boldsymbol{e}_r$, y along $\boldsymbol{e}_\theta$, and x along $-\boldsymbol{e}_\phi$. The symmetry implies electric current $\boldsymbol{j} = j\,\boldsymbol{e}_x$ where $j = cen\beta_x$. Where vacuum propagation would give $E^2 > B^2$, current is excited in the plasma to limit the growth of E^2 to $E^2 \approx B^2$. The two-fluid plasma with $\gamma > \gamma_{\rm MHD}$ can enforce $E^2 \lesssim B^2$ similar to normal MHD. The ceiling of $E^2 \approx B^2$ corresponds to the plateau $E \approx -B_{\rm bg}/2$, and j needed to sustain $E \approx -B_{\rm bg}/2$ is found from Equation (23):

$$j \approx -\frac{\partial_t (rE)_\xi}{2\pi r} \approx \frac{c}{4\pi r}\frac{d(rB_{\rm bg})}{dr} \approx -\frac{cB_{\rm bg}}{2\pi r}. \tag{82}$$

The plasma accelerated in the wave flows with speed $\beta_z \approx -1$ and density $n \approx n_{\rm bg}/(1-\beta_z) \approx n_{\rm bg}/2$ (Equation (24)), so the current j corresponds to

$$\beta_x = \frac{j}{ecn} \approx -\frac{B_{\rm bg}}{\pi e r n_{\rm bg}} = -\frac{\mu}{\pi e \mathcal{N} r}. \tag{83}$$

Note that $|\beta_x| \approx \gamma_{\rm MHD}^{-1}$. If E^2 significantly exceeded B^2, the electric field would immediately accelerate particles to a much greater β_x. This cannot happen, because the small $\beta_x \approx -\gamma_{\rm MHD}^{-1}$ already gives j sufficient to enforce the ceiling $E^2 \approx B^2$. Thus, the wave profile $E(\xi)$ in the two-fluid regime should evolve similarly to MHD—an excess of $E^2 > B^2$ will be shaved off at $E \approx -B_{\rm bg}/2$.

(2) Reducing E^2 to satisfy the ceiling of $E^2 \approx B^2$ implies a well-defined loss of electromagnetic energy of the wave and the corresponding energy gain by the plasma (Section 5.2). Energy conservation then determines the profile of γ ahead of the shock just like in the MHD regime, which gives Equation (78) for $\gamma(\xi)$.

So, we expect that the analytical expressions for γ, β_x, and $u_x = \gamma\beta_x$ derived in MHD should carry over to the unmagnetized (two-fluid) regime. The transition from MHD to the two-fluid regime occurs at $|u_x| \sim 1$. The MHD condition $|u_x| \ll 1$ may also be stated as

$$\gamma \ll \frac{1}{|\beta_x|} \approx \frac{\pi e \mathcal{N} r}{\mu} \quad \text{(MHD)}, \tag{84}$$

which is equivalent to $\gamma \ll \gamma_{\rm MHD}$ (Equation (81)). Note that $|u_x| \ll 1$ implies $\gamma \approx \gamma_z \equiv (1-\beta_z^2)^{-1/2}$, i.e., the fluid energy arises from its bulk motion along z while the internal transverse motions of $e^\pm$ are negligible. By contrast, in the unmagnetized regime the $e^\pm$ develop $|u_x| \gg 1$, i.e., the internal motions become relativistic and significantly contribute to the plasma energy. This change will affect the shock jump conditions; however, the shock speed will hardly change, since it is regulated by a different condition that leads to Equation (90) below.

It is also instructive to look at the dynamical equations describing the two-fluid $e^\pm$ motions. Taking into account the symmetry of $e^\pm$ velocities, it is sufficient to consider the positrons. The energy equation in the plateau region $E \approx -B_{\rm bg}/2$ reads

$$\frac{d\gamma}{dt} = \frac{eE\beta_x}{mc} \approx -\frac{\omega_B}{2}\,\beta_x, \tag{85}$$

where $\omega_B = eB_{\rm bg}/mc$, and d/dt is taken along the particle trajectory, so $dt = d\xi/(1-\beta_z) \approx d\xi/2$. Substituting β_x from Equation (83), one finds

$$\frac{d\gamma}{d\xi} = \frac{c\sigma_{\rm bg}}{r}, \tag{86}$$

recovering the solution given in Equation (78) for $\gamma(\xi)$.

Sustaining β_x according to Equation (83) requires a small mismatch between β_z and $\beta_{\rm D} = E/B$, as seen from the x-momentum equation,

$$\frac{du_x}{dt} = \frac{e}{mc}(E - \beta_z B) \approx \frac{\omega_B}{2}(\beta_{\rm D} - \beta_z). \tag{87}$$

Let us define $\Delta \equiv -(B+E)/B$ as a measure for the deviation of E from $-B$; $\Delta > 0$ corresponds to $E^2 > B^2$. It is easy to show that $\Delta > 1 + \beta_z \approx (2\gamma_z^2)^{-1}$ is required to achieve u_x of the unmagnetized regime, so $du_x/dt \approx -(\omega_B/2)\Delta$. The electric current j and the corresponding β_x (Equation (83)) will be sustained if $du_x/dt = \beta_x\, d\gamma/dt = -\beta_x^2\,\omega_B/2$, which requires

$$\Delta \approx \beta_x^2 \approx \left(\frac{\mu}{\pi e \mathcal{N} r}\right)^2 \approx \frac{1}{2\gamma_{\rm MHD}^2}. \tag{88}$$

The tiny positive Δ shows that E^2 slightly exceeds B^2 in the unmagnetized regime. By contrast, in the MHD regime, $\Delta \approx -(2\gamma^2)^{-1} < 0$, i.e., $E^2 < B^2$. In both cases, the ceiling of $E^2 \approx B^2$ is strongly enforced.

5.4. Shock Strength

The upstream flow is decelerated in the shock: we observed in the simulation a huge jump of the plasma proper density $q = \tilde{\rho}_{\rm d}/\tilde{\rho}_{\rm u}$, which implies a huge jump of Lorentz factor $\gamma_{\rm u}/\gamma_{\rm d} \gg 1$. By contrast, the electric field has a small jump during the peak of the monster shock: $E_{\rm d} - E_{\rm u} \approx E_{\rm d} + B_{\rm bg}/2$ is a small fraction of $E_{\rm u} \approx -B_{\rm bg}/2$. This means that the plasma immediately behind the shock still moves with a large speed: $\beta_{\rm d} = E_{\rm d}/(B_{\rm bg} + E_{\rm d}) \approx -1$ and $\gamma_{\rm d} \gg 1$. The plasma decelerates

to $\gamma < \gamma_d$ farther downstream where $E > E_d$ significantly deviates from $-B_{bg}/2$.

The ratio γ_u/γ_d is regulated by the shock jump conditions. They require fast motion of the shock relative to the downstream plasma: $\gamma_{sh|d} \approx \chi^{1/7}\sqrt{\sigma_u} \gg \sqrt{\sigma_u}$. Then, using $\gamma_{sh|d} \approx 2\gamma_{sh}\gamma_d$ (note that $\beta_d \approx -1$ is opposite to $\beta_{sh} \approx 1$), one can show that $\gamma_d \gg 1$, which corresponds to $\kappa_d \ll 1$ and $E_d + B_{bg}/2 \ll B_{bg}/2$.

The value of $\kappa_d \ll 1$ may be derived as follows. The condition that $E_d = E_0 \sin(\omega\xi_{sh})$ stays close to $-B_{bg}/2$ implies

$$\sin(\omega\xi_{sh}) \approx -\frac{B_{bg}}{2E_0} = -\frac{R_\times^2}{r^2}. \tag{89}$$

This relation controls the shock position ξ_{sh} as a function of time $t \approx r/c$ and thus determines its speed $d\xi_{sh}/dt = 1 - \beta_{sh} \approx (2\gamma_{sh}^2)^{-1}$ (when the shock Lorentz factor $\gamma_{sh} \gg 1$). This gives

$$\gamma_{sh}^2 \approx \left(2\frac{d\xi_{sh}}{dt}\right)^{-1} \approx \frac{\omega R_\times}{4c}\, x\, \sqrt{x^4 - 1}\,. \tag{90}$$

On the other hand, the shock jump conditions require $d\xi_{sh}/dt$ given in Equation (58). Substituting it into Equation (90), we find

$$\kappa_d^7 \approx \frac{\zeta^2}{32\eta}\, x^{11}(x^4 - 1), \tag{91}$$

where $\eta = (\sigma_T B_\times/8\pi e)$.

The obtained κ_u and κ_d (Equations (75) and (91)) determine $q = \tilde{\rho}_d/\tilde{\rho}_u = \kappa_d/\kappa_u$:

$$q \approx \left[\frac{4(1 - x^{-4})}{\zeta^5\, \eta\, x^{13}}\right]^{1/7}\left(\pi - 2\arcsin\frac{1}{x^2}\right). \tag{92}$$

This result is in excellent agreement with the numerical simulation (Figure 6). It loses accuracy at large radii where E_d significantly deviates from $-B_{bg}/2$, so that the approximation of Equation (89) becomes invalid. This occurs when κ_d (Equation (91)) increases to $\sim$1, i.e., at radius

$$x_1 \approx \left(\frac{32\eta}{\zeta^2}\right)^{1/15}. \tag{93}$$

Note also that $\kappa_d \approx 1$ corresponds to $q \approx 2\gamma_u$. For the wave simulated in Section 4 with $\mathcal{N} = 10^{39}$, we find $x_1 \approx 1.85$, which corresponds to $r_1 = x_1 R_\times \approx 2.6 \times 10^8$ cm.

The shock radiative parameter χ is given by Equation (56), as long as $\chi \gg 1$. Substitution of Equations (75) and (91) for κ_u and κ_d yields

$$\chi = 2^{3/2}\,\frac{\eta}{\zeta^2}\,\frac{(1 - x^{-4})^{3/4}}{x^8}\left(\pi - 2\arcsin\frac{1}{x^2}\right)^{7/2}. \tag{94}$$

It reproduces the value of χ observed in the simulation.

5.5. Precursor Effect

In an accompanying paper, we will describe an additional effect: the shock emits high-frequency precursor waves into the upstream, which can significantly decelerate the upstream flow before it reaches the collisionless shock.

The precursor emission begins immediately after shock formation at time $t_\times \approx R_\times/c$ near the coordinate $\xi_0 = 3\pi/2\omega$. As the wave expands to $r > R_\times$, the shock moves to $\xi_{sh} > \xi_0$, so the precursor occupies the growing region

$$\frac{3\pi}{2\omega} = \xi_0 < \xi < \xi_{sh}. \tag{95}$$

Note that the plasma flow in the MHD wave crosses half of the plateau $E \approx -B_{bg}/2$ ahead of the precursor, at $\xi < \xi_0$, and begins to interact with the precursor at ξ_0 with the Lorentz factor determined by Equation (78),

$$\gamma(\xi_0, r) \approx \frac{R_\times^4}{\zeta\, r^4}\left(\frac{\pi}{2} - \arcsin\frac{R_\times^2}{r^2}\right). \tag{96}$$

It is lower by a factor of 2 compared to γ_u defined previously at ξ_{sh} (Equation (75)). In the accompanying paper, we find that the precursor can decelerate the upstream flow as it moves from ξ_0 to ξ_{sh}, so the flow comes to the shock with $\gamma(\xi_{sh}) \ll \gamma(\xi_0)$ (and the shock radiative parameter χ may drop below unity). This pre-deceleration effect depends on a somewhat uncertain efficiency of precursor emission, which has not been studied in the extreme regime of the monster shocks.

Regardless of the shock structure details, with or without the precursor, the basic picture remains the same: the upstream first develops the huge Lorentz factor $\gamma(\xi_0)$ (or $2\gamma(\xi_0)$ without the precursor) and then promptly radiates its energy, before completing the shock transition. The emitted radiation is in the X-ray/gamma-ray band, as shown in Section 6 below.

5.6. Forward Shock

The plateau evolution described above leads to erasing half of the wave oscillation where $E < 0$. It does not affect the leading half $0 < \omega\xi < \pi$ where $E > 0$ (Figure 5). This leading part of the wave propagates with negligible distortions at radii $r \sim R_\times$, nearly as in vacuum. However, at larger $r \gg R_\times$ the profile of $E(\xi) > 0$ gradually steepens at the leading edge $\xi = 0$ and eventually forms a forward shock. Below we find radius R_F where this forward shock is launched.

The part of the wave with $E > 0$ has $\kappa > 1$ and satisfies $\kappa^3\sigma_{bg} \gg 1$. This implies small bending of characteristics, $d\xi_+/dt \ll 1$, and the evolution of κ along C^+ follows the simple relation $\kappa^2 = 1 + 2Kr^2$ (Appendix A).

In the zero-order of the small parameter $d\xi_+/dt$, C^+ are described by $\xi_+ = \xi_i$ or $r_+ = c(t - \xi_i)$. The correction $\xi_+ - \xi_i = (\kappa - 1)^2/8cDK^2\kappa$ (Equation (66)) in the leading order may be found iteratively by substituting $\kappa^2 = 1 + 2Kr^2$ with the zero-order $r = c(t - \xi_i)$. This gives

$$\kappa^2 = 1 + 2Kc^2(t - \xi_i)^2, \quad \kappa\frac{\partial\kappa}{\partial\xi_i}\bigg|_t = \frac{dK}{d\xi_i}r^2 - 2Krc. \tag{97}$$

One can now examine the flow of C^+ characteristics $\xi_i \to \xi_+(\xi_i, t)$ and identify where $(\partial\xi_+/\partial\xi_i)_t$ vanishes, launching a shock. For the C^+ flow with $\sigma_{bg}\kappa^3 \gg 1$, one can use Equation (97) to find

$$\frac{\partial\xi_+}{\partial\xi_i}\bigg|_t = 1 - \frac{r^3}{2D\kappa^3}\left[\frac{dK}{d\xi_i}\frac{r^3}{c}\frac{(3\kappa + 1)}{(\kappa + 1)^3} + 1\right]. \tag{98}$$

We here focus on the interval $0 < \omega\xi_i < \pi$ and observe that $\partial\xi_+/\partial\xi_i$ drops fastest at the leading edge $\xi_i = 0$, where $\kappa = 1$ is minimum and $dK/d\xi_i = K_0\omega = \omega/2R_\times^2$ is maximum. We find

at the leading edge (using $r^3 dK/d\xi_i = \omega r^3/2R_\times^2 \gg c$),

$$\left.\frac{\partial \xi_+}{\partial \xi_i}\right|_t = 1 - \frac{\omega r^6}{8cDR_\times^2} = 1 - \frac{\zeta}{8}\frac{r^6}{R_\times^6}. \quad (99)$$

Hence, the forward shock forms at radius

$$R_F = \left(\frac{8}{\zeta}\right)^{1/6} R_\times \approx 3 \times 10^9 \frac{\mu_{33}^{1/2}}{\mathcal{N}_{37}^{1/6}\nu_4^{1/6}L_{43}^{1/12}} \text{ cm}. \quad (100)$$

The wave now carries two shocks, at $\xi \approx 0$ and $\xi > 2\pi/\omega$, which slowly shift to larger ξ. The nascent forward shock at $r = R_F$ is weak. It becomes ultrarelativistic at $r \gg R_F$ and later turns into a strong blast wave expanding into the wind outside the magnetosphere.

6. Gamma Rays from Monster Shocks

First, consider shocks with neglected effects of the precursor on the upstream flow. As the plasma flow crosses the shock, it radiates its energy in curvature photons. The spectrum of curvature radiation from a particle with Lorentz factor γ cuts off exponentially at $\omega > \omega_c$ (Landau & Lifshitz 1975), where

$$\omega_c = \left(\frac{3}{2}\right)^{3/2}\left(\frac{c\,\dot{\mathcal{E}}_e}{e^2}\right)^{1/2}\gamma. \quad (101)$$

The radiated power $\dot{\mathcal{E}}_e = \dot{\gamma}mc^2$ is Lorentz invariant, and may be evaluated in the shock frame, $\dot{\gamma} = |d\gamma'/dt'|$, using the solution derived in Appendix B. It shows that γ' drops from the upstream value $\gamma_u' \approx 2\gamma_u\gamma_{sh} \approx \gamma_{sh}/\kappa_u$ with rate

$$\dot{\gamma} \approx \frac{4}{3}\frac{\gamma_u'}{\sqrt{\sigma_u}}\frac{e\tilde{B}_u}{mc}(\sqrt{g} - g^2) \approx \frac{4}{3}\frac{\gamma_{sh}\omega_B}{\sqrt{\sigma_u}}(\sqrt{g} - g^2), \quad (102)$$

where $g = \gamma'/\gamma_u'$, $\omega_B = eB_{bg}/mc$, and $\tilde{B}_u = \kappa_u B_{bg}$. The emission frequency ω_c is maximum at $g = (5/8)^{2/3}$,

$$\omega_c^{max} \approx 0.88\,\gamma_u\left(\frac{c\,\omega_B}{r_e}\frac{\gamma_{sh}}{\sqrt{\sigma_u}}\right)^{1/2}. \quad (103)$$

Substituting the solutions for $\gamma_u(r)$ (Equation (75)), $\gamma_{sh}(r)$ (Equation (90)), and $\sigma_u = \kappa_u\sigma_{bg} \approx \sigma_{bg}/2\gamma_u$, we find

$$\epsilon_c \equiv \frac{\hbar\omega_c^{max}}{mc^2} \approx \frac{1}{\zeta}\left(\frac{B_\times}{\alpha_f B_Q}\right)^{1/2} f_c(x), \quad (104)$$

where $\alpha_f = e^2/\hbar c \approx 1/137$, $B_Q = m^2c^3/\hbar e \approx 4.4 \times 10^{13}$ G, $B_\times = B_{bg}(R_\times)$, $x = r/R_\times$, and

$$f_c(x) = \frac{0.88(\pi - 2\arcsin x^{-2})^{5/4}(1 - x^{-4})^{1/8}}{2^{1/4}x^5}. \quad (105)$$

One can see that ω_c^{max} is in the far gamma-ray band.

Next, consider emission from shocks affected by the precursor (Section 5.5). Main emission now occurs where the upstream flow enters the precursor and begins deceleration with $\gamma \approx \gamma(\xi_0)$ given by Equation (96). The emission frequency is still given by Equation (101), but now with $\dot{\mathcal{E}}_e$ being the emitted power that results from the particle interaction with the precursor. Using $d\xi = (1 - \beta)dt \approx 2dt$ along the plasma streamline, one can see that $\delta\xi_{dec} \equiv 2\gamma mc^2/\dot{\mathcal{E}}_e$ defines a characteristic deceleration scale, and we rewrite Equation (101) as

$$\omega_c = \frac{3^{3/2}}{2}\gamma^{3/2}\left(\frac{c}{r_e\,\delta\xi_{dec}}\right)^{1/2}. \quad (106)$$

During the main phase of shock evolution, the upstream flow reaches the peak Lorentz factor $\gamma(\xi_0) \approx (2\zeta)^{-1}$ and radiates photons with dimensionless energy

$$\epsilon_c \approx \frac{8 \times 10^{-9}\,\nu_4^{1/2}}{(\omega\,\delta\xi_{dec})^{1/2}\zeta^{3/2}}, \quad (107)$$

where we have normalized $\delta\xi_{dec}$ to the precursor width $\xi_{sh} - \xi_0 \sim \omega^{-1}$. Note that the precursor deceleration effect is strong if $\delta\xi_{dec} \ll \omega^{-1}$, and so $\epsilon_c \gg 10^{-8}\zeta^{-3/2}$. This gives the characteristic ϵ_c in the gamma-ray band.

7. Discussion

7.1. Formation of Shocks

The magnetospheres of neutron stars have a huge magnetization parameter σ_{bg}, and therefore their low-frequency perturbations are often described as FFE waves, which propagate with the speed of light. This description is excellent near the star; however, it fails when perturbations propagate to larger radii and grow in relative amplitude E/B_{bg}. Then, FFE becomes remarkably self-destructive: it pushes itself out of the realm of its applicability $E^2 < B^2$, and the wave dynamics in the FFE limit $\sigma_{bg} \to \infty$ becomes undefined. Therefore, waves should be described using the full MHD framework, with an arbitrarily large but finite σ_{bg}.

In particular, compressive MHD waves (vacuum radio waves in the FFE limit) steepen into shocks at radius $R_\times$ (Equation (60)) for any high σ_{bg}. The higher the value of σ_{bg}, the stronger the shock at $R_\times$, as we have demonstrated by solving the MHD wave equation. As an example, we presented the evolution of a 10 kHz wave with an initial sine profile $E = E_0 \sin(\omega\xi)$, where $\xi = t - r/c$. We have also derived the wave evolution analytically to show how it depends on the wave frequency ω and power L. Our main conclusions are as follows.[8]

1. When the wave reaches $R_\times \sim 10^8$ cm, it suddenly begins to pull the background magnetosphere toward the star, creating an ultrarelativistic flow at the wave oscillation phase near $3\pi/2$ (trough), where E^2 approaches B^2 (Figure 5). The flow acceleration at the trough of the wave is accompanied by plasma expansion, reducing the local magnetization parameter σ, so the wave propagation slows down at the trough and the wave "stumbles," steepening into a shock. The accelerated plasma flow forms the upstream of the shock. It develops a huge Lorentz factor $\gamma_u \sim c\sigma_\times/\omega R_\times$, where $\sigma_\times \equiv \sigma_{bg}(R_\times)$. The peak γ_u is reached when the wave crosses $r \approx 1.15R_\times$, and here the local magnetization parameter of the flow is reduced to $\sigma_u \sim \sigma_\times/\gamma_u \sim \omega R_\times/c$, which is independent of σ_{bg}.
2. The monster shock has an unusual microstructure. The upstream flow experiences strong radiative losses before completing a single Larmor rotation, i.e., before forming

[8] These conclusions hold for kilohertz perturbations typically excited in magnetars, neutron star binaries, and neutron star collapse. The results are different for megahertz to gigahertz waves (Beloborodov 2023).

the downstream MHD flow. This affects the shock jump conditions. In particular, we found that the shock Lorentz factor relative to the downstream plasma is $\gamma_{\rm sh|d} \approx (1+\chi)^{1/7}\sqrt{\sigma_{\rm u}}$, where χ is a new dimensionless radiative parameter (Equation (55)). The shock structure is further complicated by the precursor emission, which can induce radiative losses of the upstream flow ahead of the shock; this effect will be described elsewhere.

3. As the kilohertz wave propagates beyond $R_\times$, half of its oscillation ($\pi < \omega\xi < 2\pi$) becomes erased (Figure 5). Thus, the wave loses half of its energy after crossing $R_\times$. The large wave energy per magnetospheric particle corresponds to the huge Lorentz factor gained by the plasma. The plasma radiates the gained energy in the ultrarelativistic shock, leading to a bright X-ray burst.
4. Waves accelerating the plasma to $\gamma \sim (2\omega_B/\omega)^{1/2}$ enter the two-fluid regime with unmagnetized particles. We argued in Section 5.3 that the main MHD results will carry over to the two-fluid regime. In particular, the Lorentz factor of upstream particles will be given by the same expression $\gamma \sim c\sigma_\times/\omega R_\times$, reaching enormous values in powerful low-frequency waves.
5. The leading half of the wave oscillation ($\omega\xi < \pi$) does not approach the condition $E^2 = B^2$. It crosses $R_\times$ without any significant changes. A wave consisting of only this half oscillation would not develop the monster radiative shock at $R_\times$. However, in any case, at a larger radius $R_{\rm F}$ (a few times 10^9 cm; see Equation (100)), the leading edge of the wave steepens into a forward shock.

Both $R_\times$ and $R_{\rm F}$ are well inside the typical light cylinder of a magnetar, $R_{\rm LC} \sim 10^{10}$ cm. Thus, strong kilohertz waves from magnetars inevitably launch shocks inside the magnetosphere. The monster shock forms at $\omega\xi \approx 3\pi/2$; then it weakens and shifts to $\omega\xi \approx 2\pi$. The forward shock forms at $\xi \approx 0$ and gets stronger at large radii; it will evolve into an ultrarelativistic blast wave expanding into the magnetar wind (this evolution will be further described elsewhere).

Our results clarify the relation between two pictures proposed for electromagnetic ejecta from magnetars: a vacuum-like electromagnetic wave (Lyubarsky 2014) versus a blast wave in the wind (Beloborodov 2017, 2020). Blast wave formation is important for the theory of fast radio bursts (see Lyubarsky 2021 for a review), because it produces semicoherent radio emission with submillisecond duration. Kinetic simulations demonstrating the efficiency and polarization of this emission are found in Sironi et al. (2021). In addition, Thompson (2023) recently proposed that the blast wave may produce radio waves by a different mechanism if it expands into a turbulent medium.

7.2. Compton Drag

Our calculations of magnetospheric MHD waves neglected Compton drag exerted by ambient radiation. Radiation flows from magnetars with luminosities $L_\star \sim 10^{35}$ erg s^{-1}, peaking at photon energies $\varepsilon_\star mc^2 \sim 1$ keV (Kaspi & Beloborodov 2017). When plasma in the wave develops Lorentz factor $\gamma \gg 1$, it upscatters the keV photons to energies $\varepsilon_{\rm sc} \sim \gamma^2\varepsilon_\star$ as long as $\gamma < \varepsilon_\star^{-1}$. This gives a maximum power of scattered radiation $L_{\rm sc}^{\rm max} \sim \gamma^2 L_\star \sim \varepsilon_\star^{-2} L_\star \sim 3\times 10^{40}$ erg s^{-1} (this upper bound assumes that each keV photon is upscattered). Waves with power $L \gg L_{\rm sc}$ are weakly affected by Compton drag. When $\gamma > \varepsilon_\star^{-1}$ scattering occurs with a smaller cross section $\sigma_{\rm sc} \sim \sigma_{\rm T}/\gamma\varepsilon_\star$ and $\varepsilon_{\rm sc} \sim \gamma$. Using the scattering optical depth $\tau_{\rm sc} \sim \sigma_{\rm sc}\mathcal{N}/R_\times^2$, we find

$$\frac{L_{\rm sc}}{L} \sim \frac{L_\star \varepsilon_{\rm sc}\tau_{\rm sc}}{L\,\varepsilon_\star} \sim \frac{\sqrt{8}\,\sigma_{\rm T}\mathcal{N}L_\star}{\mu(cL)^{1/2}\varepsilon_\star^2} \sim 10^{-4}\,\frac{L_{\star,35}}{\varepsilon_{\star,-3}^2}\,\frac{\mathcal{N}_{39}}{\mu_{33}L_{44}^{1/2}}. \tag{108}$$

This again gives $L_{\rm sc} \ll L$, even if the density parameter $\mathcal{N} = nr^3$ is increased by secondary $e^\pm$ creation (which mainly occurs behind the wave, not inside it; see Section 7.4 below). Thus, we conclude that the magnetar radiation does not prevent the enormous plasma acceleration in magnetosonic waves at $r \approx R_\times$.

7.3. Wave Evolution at $\theta \neq \pi/2$

In this paper, we calculated the wave propagation in the equatorial plane of a dipole magnetosphere, $\theta = \pi/2$. The obtained solution shows a clear physical picture of the wave evolution. It includes two facts that help extend the picture to $\theta \neq \pi/2$:

(1) The MHD wave is well described as a vacuum electromagnetic wave superimposed on the background dipole magnetosphere until this superposition hits the condition $E^2 = B^2$, launching the monster shock. In the equatorial plane, this condition is reached at $R_\times$. In other parts of the magnetosphere, one can find the location of shock formation from the same condition $E^2 \approx B^2$, which corresponds to drift speed $|\boldsymbol{\beta}_{\rm D}| \to 1$. The drift velocity is given by

$$\boldsymbol{\beta}_{\rm D} = \frac{\boldsymbol{E}\times\boldsymbol{B}}{B^2} = \frac{E(B_{{\rm bg},\theta} + E)\boldsymbol{e}_r - EB_{{\rm bg},r}\boldsymbol{e}_\theta}{B_{\rm bg}^2 + 2EB_{{\rm bg},\theta} + E^2}, \tag{109}$$

with $\boldsymbol{E} = -E\boldsymbol{e}_\phi$ and $\boldsymbol{B} = \boldsymbol{B}_{\rm bg} + E\boldsymbol{e}_\theta$. The wave propagates as in vacuum until $\beta_{\rm D}^2 = E^2/B^2$ approaches unity, which corresponds to $B_{\rm bg}^2 + 2EB_{{\rm bg},\theta} = 0$. This condition is first approached at the trough of the wave, $E = -E_0$. It defines the critical surface $r_\times(\theta)$:

$$E_0 = \frac{B_{\rm bg}^2}{2B_{{\rm bg},\theta}} \quad\Rightarrow\quad r_\times(\theta) = R_\times\sqrt{\frac{4-3\sin^2\theta}{\sin\theta}}. \tag{110}$$

A spherical wave with approximately isotropic power reaches this surface first at $\theta = \pi/2$, $r = R_\times$. As the wave continues its propagation to $r > R_\times$, the monster shock develops in a range of latitudes, $\sin\theta_\times < \sin\theta \leqslant 1$ (Figure 7), where $\sin\theta_\times(r)$ is found by solving Equation (110) for $\sin\theta$.

The condition $E^2 \to B^2$ corresponds to the plasma velocity $\boldsymbol{\beta}$ approaching the unit vector,[9]

$$\boldsymbol{\beta}_{\rm D} = \frac{(B_{{\rm bg},r}^2 - B_{{\rm bg},\theta}^2)\boldsymbol{e}_r + 2B_{{\rm bg},r}B_{{\rm bg},\theta}\boldsymbol{e}_\theta}{B_{\rm bg}^2} \quad (E^2 = B^2)$$
$$= \frac{(4-5\sin^2\theta)\boldsymbol{e}_r + 2\sin(2\theta)\boldsymbol{e}_\theta}{4-3\sin^2\theta}. \tag{111}$$

[9] The oblique (nonradial) plasma motion in the wave at $\theta \neq \pi/2$ develops a velocity component $\boldsymbol{\beta}_\parallel$ parallel to $\boldsymbol{B}$, as discussed in the accompanying paper (Beloborodov 2023). Then, the accurate total four-velocity is $\boldsymbol{u} = \boldsymbol{u}_\parallel + \boldsymbol{u}_{\rm D}$. However, $u_\parallel \ll u_{\rm D}$ when $E^2 \to B^2$.

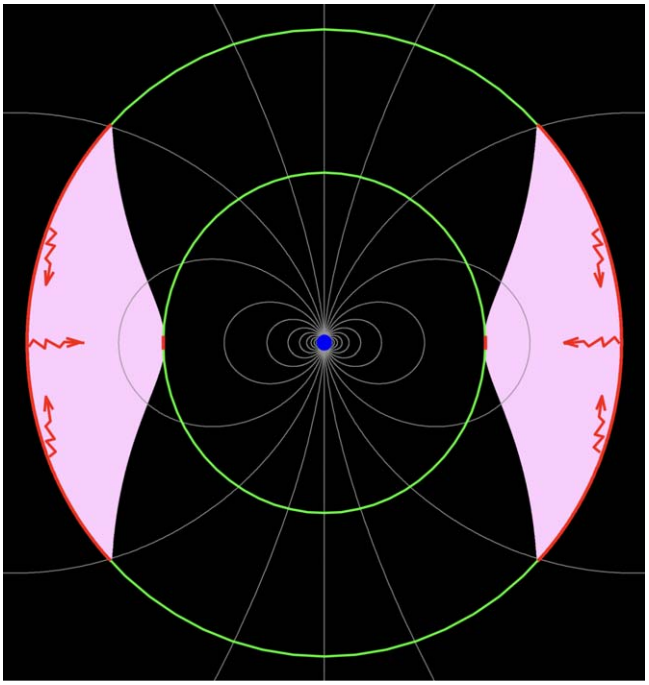

Figure 7. Expansion of a spherical wave front through the magnetosphere. The front thickness (wave duration) is much smaller than r/c. Its shock-free part ($\sin\theta < \sin\theta_\times$) is plotted in green and the part carrying the shock is in red. The region swept by the shock is shaded in pink; its boundary $\theta(r)$ is defined by Equation (110). The wave front is shown at two moments: (1) when it crosses radius $R_\times$ (the shock has formed in the equatorial plane, $\theta = \pi/2$); and (2) when it expands to a larger radius (the shock now occupies a range of θ). The shock radiates gamma rays (red arrows) beamed along $\boldsymbol{\beta}_D$ (Equation (111)); most of them convert to $e^\pm$ pairs.

This unit vector represents the direction of the accelerated plasma flow. Note that the radial component of $\boldsymbol{\beta}_D$ changes sign at $\sin\theta = 2/\sqrt{5}$.

(2) The obtained solution at $\theta = \pi/2$ demonstrates that the wave evolution outside $R_\times$ follows a simple principle: the part $\Delta\xi$ of the wave oscillation that has hit the ceiling $E^2 = B^2$ forms the plateau of $E \approx E_p$ while the rest of the oscillation profile $\boldsymbol{E}(\xi)$ follows nearly vacuum propagation (until the wave approaches $R_F \gg R_\times$). A similar description can be used at $\theta \neq \pi/2$. The plateau electric field is set by the condition $E^2 \approx B^2$:

$$\boldsymbol{E}_p(r,\theta) = \frac{B_{bg}^2}{2B_{bg,\theta}}\,\boldsymbol{e}_\phi = \frac{\mu}{2r^3}\,\frac{r_\times^2}{R_\times^2}\,\boldsymbol{e}_\phi. \tag{112}$$

It is sustained by electric current $\boldsymbol{j}_p = -\partial_r(r\boldsymbol{E}_p)/2\pi r$, as required by Equation (23). The plasma passing through the wave is accelerated on the plateau with rate $d\gamma/dt = \boldsymbol{E}_p\cdot\boldsymbol{j}_p/\rho c^2$, which gives

$$\frac{d\gamma}{d\xi} = \frac{\boldsymbol{E}_p\cdot\boldsymbol{j}_p}{\rho_{bg}c^2} = -\frac{\partial_r(r^2E_p^2)}{4\pi c\rho_{bg}r^2} = \frac{c\sigma_{bg}}{r}\,\frac{B_{bg}^2}{B_{bg,\theta}^2}, \tag{113}$$

where we used $d\xi = (1-\beta_r)dt = (\rho_{bg}/\rho)dt$ along the plasma streamline. A plateau of width $W_p = c\Delta\xi$ gives

$$\gamma \approx \sigma_{bg}\,\frac{W_p}{r}\,\frac{B_{bg}^2}{B_{bg,\theta}^2}. \tag{114}$$

The acceleration region W_p with $\boldsymbol{E} \approx \boldsymbol{E}_p$ and current $\boldsymbol{j}_p$ appear at $r = r_\times$ and grows at $r > r_\times$. For example, waves with initial profile $E(\xi) = E_0\sin(\omega\xi)$ develop plateaus with $W_p(r) \approx (c/\omega)[\pi - 2\arcsin(r_\times^2/r^2)]$.

As a result of the wave evolution at $r > r_\times$, the parts of the wave profile with $\boldsymbol{E}\cdot(\boldsymbol{k}\times\boldsymbol{B}_{bg}) < 0$ become erased and replaced by the plateau with $E_p/E_0 \to 0$ at $r \gg r_\times$. In axisymmetric waves $\boldsymbol{E}$ oscillates along $\boldsymbol{e}_\phi$, and so the parts with $E_\phi > 0$ become erased. In particular, oscillating kilohertz waves with average $\overline{E}_\phi \approx 0$ lose half of their energy. The lost energy is radiated by the shock.

7.4. Pair Creation and Emission of X-Ray Bursts

The monster shock radiates high-energy photons (Section 6), which can convert to $e^\pm$ pairs. The upstream flow entering the equatorial shock moves along $-\boldsymbol{e}_r$ with a high Lorentz factor γ, so its gamma-ray emission is beamed toward the star. These gamma rays propagate perpendicular to $\boldsymbol{B}_{bg}$ and will convert to $e^\pm$ pairs. The mean free path for conversion is given by (Erber 1966),

$$l_{ph}(r) \approx \frac{4.4\,r_e}{\alpha_f^2}\,\frac{B_Q}{B_{bg}}\exp\left(\frac{8}{3\epsilon_c}\,\frac{B_Q}{B_{bg}}\right). \tag{115}$$

For our sample models in Section 4, $l_{ph}(R_\times) \gg R_\times$ and $l_{ph}(r)$ will be reduced below r when the gamma rays propagate to $r \ll R_\times$ where $B_{bg} \gg B_\times$.

Magnetosonic waves accelerate the plasma radially only in the equatorial plane of the magnetosphere. Waves at different latitudes accelerate the plasma obliquely (Section 7.3), so the gamma-ray emission is oblique, not beamed toward the star, and can avoid absorption by the ultrastrong field. Instead, the oblique gamma rays will convert to $e^\pm$ pairs via photon-photon collisions, as they propagate toward the equatorial plane and collide with the symmetric gamma rays from the lower hemisphere. The collisions will be efficient because of the broad spectrum of curvature radiation (its half-width extends from $0.01\epsilon_c$ to $1.5\epsilon_c$; see, e.g., Longair 1994), so the gamma rays will find counterparts near the threshold for $e^\pm$ creation with a large cross section $\sim 0.1\sigma_T$. This leads to a high pair production rate, and the region behind the shock will become populated with optically thick $e^\pm$ plasma.

The created $e^\pm$ plasma will cool by emitting softer synchrotron photons and by scattering the photons as they diffuse out of the $e^\pm$ cloud behind the shock. Radiation production is controlled by the compactness parameter,

$$\ell_\times = \frac{\sigma_T L}{8\pi mc^3 R_\times} \approx 7.7\times 10^4\, L_{43}^{5/4}\mu_{33}^{-1/2}. \tag{116}$$

The high compactness implies a large optical depth to photon-photon collisions, so most gamma-rays should convert to pairs, which experience fast radiative cooling. Thus, the shock emission will be reprocessed to photons of lower energies. The luminosity of the resulting X-ray burst is comparable to the power of the original kilohertz wave. The burst resembles radiative processes in compact magnetic flares simulated in Beloborodov (2021b); its spectrum can be found with similar detailed radiative transfer simulations, which we leave for future work.

A minimum duration of the burst is set by the light-crossing time r/c, where r is a few $R_\times$. A typical value of this minimum duration is ~ 10 ms. Bursts with high $\ell_\times$ may last longer, as it takes time for the reprocessed X-rays to diffuse out of the

optically thick $e^{\pm}$ cloud. In addition, star quakes may produce multiple shocks, creating multiple bursts. Such composite bursts can last $\sim$100 ms, depending on the crustal quake coupling to the neutron star core (Bransgrove et al. 2020).

7.5. X-Ray Precursor of a Binary Merger

Consider a neutron-star binary with masses M_1 and M_2 separated by distance $2a$. The reduced mass of the binary is $M = M_1M_2/(M_1 + M_2)$, and its orbital angular velocity is

$$\Omega \approx \frac{c}{r_g}\left(\frac{a}{r_g}\right)^{-3/2}, \qquad r_g \equiv \frac{GM}{c^2}. \tag{117}$$

In a tight binary (nearing merger), Ω is in the kilohertz band.

Suppose both stars are strongly magnetized. An interface between their magnetospheres forms where their pressures balance. For instance, consider two stars with equal magnetic dipole moments $|\boldsymbol{\mu}_1| = |\boldsymbol{\mu}_2| = \mu$. Then, the interface is in the middle, at the distance a from each star, and the magnetic field at the interface is $B_i \approx \mu/a^3$.

If the orbiting stars are not in synchronous rotation, there is differential rotation between the two magnetospheres with an angular frequency $\Omega_{\rm diff}$ comparable to the orbital Ω. The two magnetic moments $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ in general are not aligned, and then the magnetic pressure at the interface will oscillate with frequency $\Omega_{\rm diff}$. The pressure variation timescale $\sim\Omega_{\rm diff}^{-1}$ exceeds the Alfvén crossing time (a few a/c), which reduces the efficiency of low-frequency wave excitation. However, the interface is also a source of higher-frequency waves, because it is prone to instabilities of the Kelvin–Helmholtz/diocotron type. The instability will generate vortices of sizes up to a fraction of a (with frequencies $\omega > c/a$), creating traction between the rotating magnetospheres.

The turbulent region around the interface will emit both Alfvén and magnetosonic waves, and further investigation of this process requires numerical simulations. An upper bound on the power deposited into the turbulence may be roughly estimated as $L_{\rm max} \sim \Omega_{\rm diff}a^3B_i^2/8\pi$. The power of magnetosonic wave emission $L \sim ca^2(\delta B)^2/8\pi < L_{\rm max}$ implies an upper limit on the emission amplitude,

$$b \equiv \frac{\delta B}{B} < \left(\frac{\Omega_{\rm diff}R}{c}\right)^{1/2} \sim \left(\frac{r_g}{a}\right)^{1/4}\left(\frac{\Omega_{\rm diff}}{\Omega}\right)^{1/2}. \tag{118}$$

In a more general case, the magnetic dipole moments of the two stars are not equal, $\mu_1 < \mu_2$. Then the interface will form at a distance R from M_1, defined by

$$\frac{\mu_1}{R^3} \approx \frac{\mu_2}{(2a - R)^3}. \tag{119}$$

The wave source will have a size of $\sim R < a$, and its power may be estimated as

$$L \sim cR^2\,\frac{b^2B_i^2}{8\pi}. \tag{120}$$

The emitted magnetosonic waves will steepen into shocks at radius $R_\times \sim (2b)^{-1/2}R$ (if it is inside the light cylinder of the binary, $R_\times < c/\Omega$).

The emission of kilohertz waves accompanied by shocks and turbulent dissipation will continue as long as the two stars orbit each other, until they eventually merge. The binary loses orbital energy to gravitational wave emission and shrinks on the timescale

$$t \sim \frac{r_g}{c}\left(\frac{a}{r_g}\right)^4, \tag{121}$$

which is also comparable to the time remaining until the merger. The wave energy generated during the remaining time t may be estimated as tL, and we obtain

$$tL \sim \frac{b^2\mu_1^2}{8\pi r_g^3}\left[\frac{1}{2}\left(1 + \frac{\mu_1^{1/3}}{\mu_2^{1/3}}\right)\right]^4. \tag{122}$$

A significant energy fraction of the MHD perturbations will be dissipated and converted to X-rays. This will occur mainly through the local turbulent dissipation near the interface and may also launch shocks at larger distances. The dissipation of perturbations will produce a detectable X-ray precursor of the merger if the stars have strong magnetic fields. For instance, a binary with $\mu_1 = \mu_2 = 10^{32}$ G cm^3 may generate X-ray luminosity $L_X \sim 10^{46}$ erg s^{-1} at $t \sim 10$ s before the merger. Wave dissipation provides an alternative to the recently proposed precursor emission by magnetic flares (Most & Philippov 2020; Beloborodov 2021b). The latter mechanism was shown to operate in binaries with antialigned $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$; it is driven by the over-twisting of magnetic flux tubes connecting the two stars.

7.6. Shocks from Neutron Star Collapse

A massive neutron star can be born (e.g., in a merger) with fast rotation, which temporarily supports it against collapse. Such objects are likely strongly magnetized and gradually lose rotation by emitting angular momentum in a magnetized wind. The spindown can eventually lead to the collapse of the massive neutron star into a black hole, producing an electromagnetic transient (Lehner et al. 2012; Falcke & Rezzolla 2014; Most et al. 2018). Numerical simulations of the collapse show that it launches a strong outgoing electromagnetic pulse propagating through the outer dipole magnetosphere. A monster shock will form where the pulse amplitude becomes comparable to the dipole field of the pre-collapse magnetosphere and $B^2 - E^2$ reaches zero. This condition was indeed seen in vacuum and FFE simulations of the magnetosphere of a collapsing star (see Figure 10 in Lehner et al. 2012). Demonstrating shock formation and tracking its propagation requires a full MHD calculation, as shown in the present paper. The shock will dissipate a significant fraction of the outgoing electromagnetic pulse and produce a bright X-ray transient.

Acknowledgments

The author is grateful to Alex Chen for sharing particle-in-cell simulations of strong magnetosonic waves prior to publication. This work was prompted, in large part, by those results, which demonstrate shock formation through first principles, fully kinetic calculations. The author is also grateful to the anonymous referee for providing many useful comments, to Yuri Levin and Ashley Bransgrove for discussions of fast magnetosonic waves from quakes, and to Jens Mahlmann, Emanule Sobacchi, and Sasha Philippov for reading the manuscript and providing comments. This work is supported by NSF AST 2009453, NASA 21-ATP21-0056 and Simons Foundation No. 446228.

Appendix A
Shock Formation in a Cold Magnetosonic Wave

Below, we describe analytically the flow of C^+ characteristics. We focus here on the part of C^+ flow with $E<0$ (or, equivalently, $\kappa<1$), which experiences strong deformation, leading to caustic formation.

A.1. Evolution of the Plasma Compression Factor κ *along* C^+

The relation between κ and r along C^+ is found from the ratio in Equations (45) and (43),

$$\frac{dr}{d\kappa}=\frac{r(4\sigma_{\rm bg}\kappa^3-1)}{4\sigma_{\rm bg}\kappa^2(\kappa^2-1)}\quad\Rightarrow\quad\frac{d\sigma_{\rm bg}}{d\kappa}=\frac{3\sigma_{\rm bg}\kappa^3-3/4}{\kappa^2(1-\kappa^2)}\qquad(\text{along } C^+),\tag{A1}$$

where we used $d\ln\sigma_{\rm bg}/d\ln r=-3$. This is a linear differential equation for $\sigma_{\rm bg}(\kappa)$. Its integrating factor is $(1-\kappa^2)^{3/2}$, and its solution is

$$\sigma_{\rm bg}(\kappa)=\frac{3}{8(1-\kappa^2)^{3/2}}\left[\frac{8}{3}\sigma_\star+\frac{2}{\kappa}\sqrt{1-\kappa^2}+4\arctan\left(\frac{1-\sqrt{1-\kappa^2}}{\kappa}\right)-\pi\right],\tag{A2}$$

where $\sigma_\star$ is a constant (defined for each characteristic). Note that

$$\sigma_\star=\lim_{\kappa\to 1}[(1-\kappa^2)^{3/2}\sigma_{\rm bg}]\tag{A3}$$

is set at small radii where $\sigma_{\rm bg}\gg\sigma_\star$. In this inner zone, $|\kappa-1|\ll 1$ corresponds to slow plasma motions in the wave, $|\beta|\ll 1$, which is equivalent to $B_{\rm bg}\gg|E|$. Comparing the behavior of $\sigma_{\rm bg}=\sigma_\star(1-\kappa^2)^{-3/2}$ with $\sigma_{\rm bg}=D/r^3$ (Equation (8)) one can see that

$$1-\kappa^2=\frac{r^2}{r_\star^2},\qquad r_\star=\left(\frac{D}{\sigma_\star}\right)^{1/3}.\tag{A4}$$

For each C^+, the regime $|\kappa-1|\ll 1$ holds at radii $r\ll r_\star$. In this zone, the radial position of C^+ follows vacuum propagation: $r=r_{\rm vac}=c(t-\xi_{\rm i})$. Note that $r_\star(\xi_{\rm i})$ is different for different characteristics.

Equation (A4) is equivalent to Equation (48), and so the constants $r_\star$ and K are related by

$$r_\star(\xi_{\rm i})=\frac{1}{\sqrt{-2K(\xi_{\rm i})}}.\tag{A5}$$

Equation (A4) holds as long as $1-\beta_+\ll 1$, which corresponds to $\kappa\gg\sigma_\star^{-1/3}$. We are interested in a more general regime, where β_+ can significantly drop or even change sign (which means that the C^+ characteristic turns back to the star). This happens if κ decreases below $(4\sigma_\star)^{-1/3}$. In this extreme case, the variation of r with κ becomes nonmonotonic; it occurs on a small scale $\delta r\ll r$ and can be seen only when retaining the small terms in Equation (A2).

Let us now find the solution for $\kappa(t)$ along C^+. Using Equation (A2) in Equation (43), we obtain

$$\left.\frac{d\kappa}{dt}\right|_{C^+}=-\frac{c\sqrt{1-\kappa^2}}{\kappa\, r_\star}\left[1+\frac{(1-\kappa^2)^{3/2}}{4\sigma_\star\kappa^3}\right]^{-1}\left[1+\mathcal{O}\left(\frac{1}{\kappa\sigma_\star}\right)\right].\tag{A6}$$

Note that κ monotonically decreases with t. This equation can be integrated as follows:

$$\sqrt{1-\kappa^2}+\frac{(1-\kappa^2)^2}{4\sigma_\star\kappa}=\frac{c(t-\xi_{\rm i})}{r_\star}=\frac{r_{\rm vac}}{r_\star}.\tag{A7}$$

The second term on the left-hand side (LHS) makes a negligible correction to $\sqrt{1-\kappa^2}$ if $4\sigma_\star\kappa^3\gg 1$, which corresponds to $1-\beta_+\ll 1$. Then, $\kappa=\sqrt{1-r_{\rm vac}^2/r_\star^2}=\sqrt{1+2Kr_{\rm vac}^2}$. If the condition $4\sigma_\star\kappa^3\gg 1$ is not satisfied, one must keep both terms on the LHS. This may happen only at $\kappa\ll 1$, and hence the deviation of $\kappa(t)$ from the solution $\kappa=\sqrt{1-r_{\rm vac}^2/r_\star^2}$ can be found using expansion in $\kappa\ll 1$. This expansion simplifies Equation (A7) to a depressed cubic equation,

$$\kappa^3+2\left(\frac{r_{\rm vac}}{r_\star}-1\right)\kappa-\frac{1}{2\sigma_\star}=0\qquad\left(\frac{1}{\sigma_\star}\ll\kappa\ll 1\right).\tag{A8}$$

It has one real root if $\sigma_\star^{2/3}(1-r_{\rm vac}/r_\star)<3/2^{7/3}$. Otherwise, it has three real roots, but only one of them is relevant—the root branch that gives $\kappa>0$ and approaches $\kappa=\sqrt{1-r_{\rm vac}^2/r_\star^2}$ when $\sigma_\star^{2/3}(1-r_{\rm vac}/r_\star)\gg 1$.

The evolution of κ along C^+ with $K<0$ may now be summarized as follows:

$$\kappa(t)=\left\{\begin{array}{cc}\sqrt{1-\frac{r_{\rm vac}^2}{r_\star^2}} & y\gg 1\\ 2\sqrt{y}\,\sigma_\star^{-1/3}\cos\left[\frac{1}{3}\arccos\left(\frac{1}{4y^{3/2}}\right)\right] & 2^{-4/3}<y\ll\sigma_\star^{2/3}\\ 2\sqrt{y}\,\sigma_\star^{-1/3}\cosh\left[\frac{1}{3}{\rm arccosh}\left(\frac{1}{4y^{3/2}}\right)\right] & 0<y<2^{-4/3}\\ -2\sqrt{-y}\,\sigma_\star^{-1/3}\sinh\left[\frac{1}{3}{\rm arcsinh}\left(-\frac{1}{4\,|y|^{3/2}}\right)\right] & y<0\end{array}\right\},\quad \begin{array}{c}y\equiv\frac{2}{3}\,\sigma_\star^{2/3}\left(1-\frac{r_{\rm vac}}{r_\star}\right)\\ r_{\rm vac}\equiv c(t-\xi_{\rm i}).\end{array}\tag{A9}$$

These real expressions fully describe $\kappa(t)$ along C^+ (note that $\sigma_\star\gg 1$ implies a large overlap of $y\gg 1$ and $y\ll\sigma_\star^{2/3}$).

A.2. Crossing of Characteristics

The crossing of C^+ characteristics is determined by the solution for their shapes $\xi_+(t)$. It is easier to solve for $\xi_+(\kappa)$ first, and then use the (monotonic) relation between κ and t along C^+ (Appendix A.1). The solution for $\xi_+(\kappa)$ can be found from the ratio of Equations (44) and (43),

$$\frac{d\xi_+}{d\kappa}=\frac{r}{2c\,\sigma_{\rm bg}\kappa^2(\kappa^2-1)}.\tag{A10}$$

We substitute here $\sigma_{\rm bg}(\kappa)$ (Equation (A2)) and $r=(\sigma_\star/\sigma_{\rm bg})^{1/3}r_\star$ to obtain a closed differential equation for $\xi_+(\kappa)$. Its integration (with the initial condition $\xi_+=\xi_{\rm i}$ at $\kappa=1$) yields

$$\xi_+=\xi_{\rm i}+\frac{r_\star}{2c\,\sigma_\star\kappa}\left[(1-\kappa)^2+\mathcal{O}\left(\frac{1}{\sigma_\star\kappa}\right)\right].\tag{A11}$$

One can think of the initial value $\xi_{\rm i}$ as a Lagrangian coordinate in the flow of C^+ characteristics and quantify deformation of the wave profile using "strain" $\partial\xi_+/\partial\xi_{\rm i}$ evaluated at $t=$ const. A reduction of $\partial\xi_+/\partial\xi_{\rm i}$ below unity means compression of the profile, i.e., steepening of the wave. The caustic appears where $\partial\xi_+/\partial\xi_{\rm i}$ vanishes. This will occur in the region of $\sigma_\star^{-1}\ll\kappa\ll 1$, and in this region the expression for $\partial\xi_+/\partial\xi_{\rm i}$ simplifies to

$$\left.\frac{\partial\xi_+}{\partial\xi_{\rm i}}\right|_t\approx 1-\frac{2(1+\dot r_\star)}{4\sigma_\star\kappa^3+1},\quad\text{where}\quad \dot r_\star\equiv\frac{dr_\star}{d\xi_{\rm i}}=r_\star^3\dot K,\qquad \dot K\equiv\frac{dK}{d\xi_{\rm i}}.\tag{A12}$$

Here, we used Equation (A8) to express $(\partial\kappa/\partial\xi_+)_t$ and retained only the leading terms of the expansion in $\kappa\ll 1$. Thus, for each characteristic C^+ with $K<0$, we find that $\partial\xi_+/\partial\xi_{\rm i}$ vanishes when $\kappa=\kappa_{\rm v}$, where $\kappa_{\rm v}$ satisfies the relation

$$\kappa_{\rm v}^3(\xi_{\rm i})\approx\frac{2\dot r_\star+1}{4\sigma_\star}\qquad(\kappa_{\rm v}\ll 1).\tag{A13}$$

The condition $\kappa=\kappa_{\rm v}$ corresponds to time $t_{\rm v}$ determined by Equation (A8):

$$\frac{t}{r_\star}=1-\frac{\kappa^2}{2}+\mathcal{O}(\kappa^4)+\frac{1+\mathcal{O}(\kappa)}{4\sigma_\star\kappa}\quad\Rightarrow\quad t_{\rm v}(\xi_{\rm i})\approx\frac{r_\star}{c}\left(1-\frac{\kappa_{\rm v}^2}{2}+\frac{1}{4\sigma_\star\kappa_{\rm v}}\right).\tag{A14}$$

The small difference $t_{\rm v}-r_\star/c$ is determined by the small $\kappa_{\rm v}\ll 1$, which is given in the leading order by Equation (A13).

The caustic forms on the characteristic C^+ for which $\partial\xi_+/\partial\xi_{\rm i}=0$ is reached first, i.e., where $t_{\rm v}(\xi_{\rm i})$ is minimum. This characteristic (labeled as $\xi_{\rm i}^{\rm c}$) satisfies the condition $dt_{\rm v}/d\xi_{\rm i}=0$. Using Equation (A13), we find

$$\frac{d\kappa_{\rm v}}{d\xi_{\rm i}}=\frac{\ddot r_\star}{6\sigma_\star\kappa_{\rm v}^2}+\frac{\kappa_{\rm v}\dot r_\star}{r_\star},\quad\text{where}\quad \ddot r_\star=\frac{3\dot r_\star^2}{r_\star}-\frac{r_\star\ddot K}{2K},\tag{A15}$$

and then obtain the condition $dt_v/d\xi_i = 0$ in the form

$$\frac{dt_v}{d\xi_i} = (1 + \dot{r}_\star)\left(1 + \frac{r_\star^2 \ddot{K}}{24\sigma_\star^2 \kappa_v^2 K}\right) + Z = 0. \tag{A16}$$

The terms collected in $Z(\kappa_v)$ are small, and we find a simple expression for the fluid compression factor at the caustic, $\kappa_c \equiv \kappa_v(\xi_i^c)$ (one can show that $Z(\kappa_c) \ll 1 + \dot{r}_\star$):

$$\kappa_c^4 \approx -\frac{r_\star^2 \ddot{K}}{24\sigma_\star^2 K}. \tag{A17}$$

For example, consider waves with a harmonic profile $K(\xi_i)$. Let us define $\psi \equiv \omega\xi_i$ so that $K(\psi) = K_0 \sin\psi$. Then,

$$r_\star = \frac{R_\times}{\sqrt{-\sin\psi}}, \qquad \sigma_\star = \sigma_\times(-\sin\psi)^{3/2}, \qquad \dot{r}_\star = \frac{R_\times \omega \cos\psi}{2c(-\sin\psi)^{3/2}}, \qquad \frac{\ddot{K}}{K} = -\omega^2. \tag{A18}$$

The caustic develops at a Lagrangian coordinate $\psi_c = \omega\xi_i^c$, and Equations (A13) and (A17) give

$$4\,(-\sin\psi_c)^{3/2}\,\kappa_c^3 = \frac{\zeta\cos\psi_c}{(-\sin\psi_c)^{3/2}} + \frac{1}{\sigma_\times}, \qquad \kappa_c^4 = \frac{\zeta^2}{24\sin^4\psi_c}, \quad \text{where} \quad \zeta \equiv \frac{\omega R_\times}{c\sigma_\times}. \tag{A19}$$

These equations hold if $\kappa_c \ll 1$, which requires $\zeta \ll 1$. We obtain from these equations,

$$\cos\psi_c = \frac{4\sqrt{\zeta}}{24^{3/4}} - \frac{c}{\omega R_\times}(-\sin\psi_c)^{3/2}. \tag{A20}$$

We consider only short waves, $\omega R_\times/c \gg 1$. Equation (A20) then implies $|\cos\psi_c| \ll 1$, and so in the leading order, one can set $\sin\psi_c \approx -1$ in Equations (A19) and (A20). This gives the relations stated in Equation (67).

The caustic forms at time $t_c = t_v(\xi_i^c)$ given by Equation (A14). Expanding $r_\star/R_\times = (1 - \cos^2\psi_c)^{-1/4} \approx 1 + (1/4)\cos^2\psi_c$ and using Equation (67), we obtain Equation (68).

The caustic position in the wave, ξ_c, may be found using the relation between ξ_+ and κ along C^+ (Equation (A11)). It gives $\xi_c \approx \xi_i^c + r_\star/2c\sigma_\star\kappa_c$, which leads to Equation (69).

Appendix B
Shock Microstructure

The flow is composed of two streams, e^+ and e^-, which have different (symmetric) trajectories. The $e^\pm$ flow is neutral everywhere, i.e., its net charge density is zero. As the flow enters the shock, it develops a transverse electric current $\boldsymbol{j}$, created by the opposite $e^\pm$ motions along $\boldsymbol{E}$. This current controls the self-consistent change of the magnetic field across the shock. In this appendix t, $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{\beta}$, u^α, and n will denote quantities measured in the shock rest frame (elsewhere in the paper, this notation is used for the lab-frame quantities).

We will use local Cartesian coordinates x, y, z with the x-axis along $\boldsymbol{E}$ and the z-axis along the upstream flow. Then, the magnetic field $\boldsymbol{B}$ is along y. The upstream quantities ahead of the shock will be denoted with subscript "u." The upstream flow will be approximated as cold, i.e., the $e^\pm$ are at rest in the plasma drift frame (the frame $\tilde{\mathcal{K}}$ where the electric field vanishes, $\tilde{E} = 0$). In this Appendix, we use units of $c = 1$.

The steady flow is described by quantities that depend only on z. In this approximation, all time derivatives vanish, $\partial_t = 0$, and hence $\nabla \times \boldsymbol{E} = -\partial_t \boldsymbol{B} = 0$. In addition, $\nabla \cdot \boldsymbol{E} = 0$ as the net charge density is zero. This implies

$$\boldsymbol{E}(z) = const = \boldsymbol{E}_u = (E_u, 0, 0). \tag{B1}$$

The $e^\pm$ streams ahead of the shock have equal velocities

$$\boldsymbol{\beta}_u^\pm = (0, 0, \beta_u), \qquad \beta_u = \frac{E_u}{B_u}. \tag{B2}$$

The straight inertial motion of $e^\pm$ describes the cold upstream flow with no gyration in the drift frame $\tilde{\mathcal{K}}$; the $e^\pm$ experience zero force $\pm e(\boldsymbol{E}_u + \boldsymbol{\beta}_u^\pm \times \boldsymbol{B}_u) = 0$.

B.1. Shock Structure Equations

As the plasma crosses the shock, the $e^{\pm}$ four-velocities $u_{\pm}^{\alpha} = (\gamma_{\pm}, \boldsymbol{u}_{\pm})$ and B change from their upstream values. The profiles $u_{\pm}^{\alpha}(z)$ and $B(z)$ obey the equations:

$$m\,\beta_z \frac{d\gamma_{\pm}}{dz} = \pm e\boldsymbol{E}\cdot\boldsymbol{\beta}_{\pm} - \dot{\mathcal{E}}_e, \qquad m\,\beta_z \frac{d\boldsymbol{u}_{\pm}}{dz} = \pm e(\boldsymbol{E} + \boldsymbol{\beta}_{\pm}\times\boldsymbol{B}) - \boldsymbol{f}_{\pm}, \qquad \nabla\times\boldsymbol{B} = 4\pi\boldsymbol{j}, \tag{B3}$$

where $\dot{\mathcal{E}}_e$ is the power radiated by the ultrarelativistic particle, and $\boldsymbol{f}_{\pm} = \boldsymbol{\beta}_{\pm}\dot{\mathcal{E}}_e$ is the radiated momentum. Similar equations, but without radiation reaction, were previously used to describe magnetosonic solitons, which have no dissipation (Kennel & Pellat 1976; Alsop & Arons 1988).

The particle entering the shock radiates because it experiences acceleration $(\pm e/m)(\boldsymbol{E} + \boldsymbol{\beta}_{\pm}\times\boldsymbol{B})$. The power radiated by a particle with four-velocity $(\gamma, \boldsymbol{u})$ is given by (e.g., Landau & Lifshitz 1975),

$$\dot{\mathcal{E}}_e = \frac{\sigma_{\rm T}}{4\pi}[(\gamma\boldsymbol{E} + \boldsymbol{u}\times\boldsymbol{B})^2 - (\boldsymbol{u}\cdot\boldsymbol{E})^2] = \frac{\sigma_{\rm T}}{4\pi}[(\gamma B - u_z E)^2 + E^2 - B^2], \tag{B4}$$

where $\sigma_{\rm T} = 8\pi r_e^2/3$ is the Thomson cross section.

Equations (B3) have a symmetry: the $e^{\pm}$ develop $\beta_x^- = -\beta_x^+$, $\beta_z^- = \beta_z^+$, and keep $\beta_y^- = \beta_y^+ = 0$. Therefore, it is sufficient to solve for the four-velocity of e^+. Below, this will be denoted by u^{α}, omitting the subscript "+." The symmetry also implies that the electric current has the form $\boldsymbol{j} = (j, 0, 0)$, where $j = en\beta_x$, $\boldsymbol{\beta} \equiv \boldsymbol{\beta}_+$ and $n = n_+ + n_-$ is the local plasma density. The current is related to the uniform particle flux $F = n\beta_z = n_{\rm u}\beta_{\rm u}$,

$$j = eF\,\frac{\beta_x}{\beta_z} = en_{\rm u}\beta_{\rm u}\,\frac{\beta_x}{\beta_z}. \tag{B5}$$

The current j is expected to be negative, since B is compressed in the shock: $dB/dz = -4\pi j > 0$. This requires $\beta_x < 0$. The solution presented below indeed shows that the e^+ flow entering the shock develops negative acceleration along x: $\beta_z du_x/dz = (e/m)(E - \beta_z B) < 0$. This is consistent with energy conservation: the compression of B occurs at the expense of the flow kinetic energy, and hence $\boldsymbol{E}\cdot\boldsymbol{j} < 0$ (energy is extracted from the plasma).

Note that γ and u_z determine all components of four-velocity u^{α} (u_x is found from $u^{\alpha}u_{\alpha} = -1$ using $u_y = 0$), and hence also determine β_x and β_z. As a result, we have three coupled equations for three unknowns γ, u_z, and B. The flow solution has three parameters: $n_{\rm u}$, $B_{\rm u}$, and $\beta_{\rm u}$. Given these parameters, the system of Equations (B3) can be integrated numerically. As shown below, the solution also admits a simple analytical approximation.

A normal collisionless shock with no radiation reaction ($\chi = 0$) would be described by two dimensionless parameters: $\gamma_{\rm u} = (1 - \beta_{\rm u}^2)^{-1/2}$ and $\sigma_{\rm u} = B_{\rm u}^2/4\pi m n_{\rm u}\gamma_{\rm u}$. Its downstream moves with Lorentz factor $\gamma_{\rm d} \approx \sqrt{\sigma_{\rm u}}$ relative to the shock. The shock structure is normally presented using fields normalized to $B_{\rm u}$ and distance normalized to the Larmor scale $r_{\rm L} \equiv m\gamma_{\rm u}/eB_{\rm u}$. By contrast, the presence of radiation reaction in Equations (B3) implies that $B_{\rm u}$ cannot be removed by the normalization of fields, since $\dot{\mathcal{E}}_e$ is quadratic in the fields. An additional scale $r_e = e^2/m$ appears in the problem, and the value of $B_{\rm u}$ controls the ratio $r_e/r_{\rm L}$. To disentangle this problem, it helps to introduce new variables:

$$g \equiv \frac{\gamma}{\gamma_{\rm u}}, \quad w \equiv \frac{\sigma_{\rm u}(\gamma - u_z)}{\gamma_{\rm u}}, \quad b \equiv \sigma_{\rm u}\left(\frac{B}{B_{\rm u}} - 1\right). \tag{B6}$$

Instead of coordinate z, we will use dimensionless s defined by

$$ds = \frac{eB_{\rm u}}{m\gamma_{\rm u}\sqrt{\sigma_{\rm u}}}\,\frac{dz}{\beta_z}. \tag{B7}$$

Equations (B3) then become

$$\frac{dg}{ds} = -\beta_{\rm u}\mathcal{F} - \chi\mathcal{R}, \qquad \frac{dw}{ds} = -\mathcal{F}\left(b + \frac{\epsilon}{2}\right) - w\chi\mathcal{R}, \qquad \frac{db}{ds} = \beta_{\rm u}\mathcal{F}, \tag{B8}$$

where

$$\epsilon \equiv \frac{\sigma_{\rm u}}{\gamma_{\rm u}^2}, \qquad \chi \equiv \frac{\sigma_{\rm T}B_{\rm u}\gamma_{\rm u}^2}{4\pi e\,\sigma_{\rm u}^{3/2}} = \frac{2}{3\epsilon^{3/2}}\frac{r_e}{r_{\rm L}}, \tag{B9}$$

$$\mathcal{F}(g, w, b) \equiv \frac{1}{g}\sqrt{w\left(2g - \frac{w}{\sigma_{\rm u}}\right) - \epsilon} = -\beta_x\sqrt{\sigma_{\rm u}}, \tag{B10}$$

$$\mathcal{R}(g, w, b) \equiv \left(w + \frac{g - w/\sigma_u}{1 + \beta_u}\epsilon\right)^2 - \epsilon^2 + 2bgw + \frac{2bg(g - w/\sigma_u)}{1 + \beta_u}\epsilon - 2b\epsilon + g^2b^2. \tag{B11}$$

Note that the presented form of the flow Equations (B8) uses β_u, ϵ, and χ as the three parameters of the shock instead of γ_u, σ_u, and B_u. The parameter σ_u appears only in the functions $\mathcal{F}$ and $\mathcal{R}$, where it enters through the term $g - w/\sigma_u$. We will verify below that $w/\sigma_u \ll g$ throughout the shock transition, and so σ_u drops out.

The choice of β_u, ϵ, and χ as the independent parameters significantly simplifies the problem. For strong ultrarelativistic shocks of interest, (i) one can set $\beta_u = 1$, and (ii) ϵ turns out so small that it effectively drops out. As a result, χ remains the only important parameter (while σ_u and γ_u enter through the definition of variables g, w, b).

In the limit of $\chi \to 0$, the autonomous system of Equations (B8) admits an immediate analytical solution. Its first integrals $g(b)$ and $w(b)$ are

$$g = 1 - b, \qquad w = \frac{b(b + \epsilon) + \epsilon\beta_u}{2\beta_u} \qquad (\chi = 0), \tag{B12}$$

where we took into account that the solution starts in the upstream with $b = 0$, $g = 1$, $w = \epsilon/2$. The first integrals correspond to the fluxes of specific energy and momentum staying uniform along z, as losses are neglected.

The monster shocks discussed in this paper propagate in the qualitatively different regime of $\chi \gg 1$. The value of χ is controlled by the shock strength developing in the kilohertz wave, as discussed in Section 5.

B.2. Boundary Condition

In order to integrate Equation (B8), one needs boundary conditions $g = g_0$, $w = w_0$, $b = b_0$ at some s_0. We will set the boundary conditions in the upstream, ahead of the shock. Far into the upstream, the flow is unaware of the shock and has $g = 1$, $w = \epsilon/2$, $b = 0$, which corresponds to $\gamma = \gamma_u$, $u_z = \gamma_u\beta_u$, $B = B_u$. However, one cannot impose these simple boundary conditions, because Equations (B8) have a trivial uniform solution with these values. The shock solution must have an infinitesimal deviation from the uniform solution at $s_0 \to -\infty$. It is obtained by choosing g_0, w_0, and b_0 slightly perturbed from 1, $\epsilon/2$, and 0.

The upstream flow with g, w, $b = 1$, $\epsilon/2$, 0 has $\mathcal{R} = 0$, and the deviation of g, w, b from 1, $\epsilon/2$, 0 begins with negligible losses ($\chi\mathcal{R} \ll \mathcal{F}$). Hence, one can choose s_0 where the flow still satisfies the adiabatic solution of Equation (B12). Only one of the three perturbed quantities at s_0 is independent: by choosing a small perturbation $b_0 \neq 0$, we also determine g_0 and w_0 according to Equation (B12). Furthermore, changing b_0 is equivalent to shifting s_0 in the profile of $b(s)$. This shift is irrelevant, since the problem has translational symmetry along s. Thus, no new free parameters are introduced by the choice of boundary conditions. The shock transition will follow a unique solution apart from the arbitrary shift along the s-axis.

In practice, b_0 does not need to be infinitesimal. It is sufficient to choose any finite small b_0 in the zone where the adiabatic solution of Equation (B12) still holds, i.e., where $\dot{\mathcal{E}}_e \ll e|\boldsymbol{E} \cdot \boldsymbol{\beta}|$. This requires

$$\chi\mathcal{R}(g_0, w_0, b_0) \ll \mathcal{F}(g_0, w_0, b_0). \tag{B13}$$

For shocks of main interest ($\beta_u \to 1$ and small ϵ, σ_u^{-1}), we find, using Equation (B12), that

$$\mathcal{R} \approx b^2\left(1 - \frac{b}{2}\right)^2, \quad \mathcal{F} \approx \frac{b}{\sqrt{1 - b}} \quad (\dot{\mathcal{E}}_e \ll e|\boldsymbol{E} \cdot \boldsymbol{\beta}|). \tag{B14}$$

The adiabatic condition (B13) is satisfied if $b_0 \ll \chi^{-1}$. In the numerical models presented below, we set $b_0 = \epsilon^{1/2}$. This satisfies the condition (B13) for shocks with $\epsilon \ll \chi^{-2}$.

One can also see here that shocks with $\chi \ll 1$ have negligible losses ($\chi\mathcal{R} \ll \mathcal{F}$) throughout the shock transition, as g drops from unity and $b = 1 - g$ changes from 0 to $b \sim 1$. By contrast, $\chi \gg 1$ implies almost immediate onset of radiation reaction as the flow enters the shock.

B.3. Numerical Solution

Numerical solutions for shocks with different values of χ are shown in Figure 8. They are obtained in the limit of $\sigma_u \gg 1$ and $\epsilon \ll 1$ (we used $\sigma_u = 10^3$ and $\epsilon = 10^{-16}$ when integrating Equations (B8); however, the result is independent of these choices). Note that the flow keeps $\beta_z \approx 1$ everywhere, and hence $s \propto z$. One can see that in the shock the normalized Lorentz factor of the flow $g(s) = \gamma/\gamma_u$ decreases while the magnetic field becomes compressed: $b(s)$ grows from zero up to the final downstream value. The relation between g and b across the shock transition is shown in Figure 8.

The presented solutions end where the flow develops gyration in the drift frame $\tilde{K}$. Beyond this point, the maser instability is expected to destroy the coherent gyration of the $e^{\pm}$ streams (Langdon et al. 1988). This instability is not followed by our calculation, and this is not needed to see the effect of radiation reaction on the shock structure. After the development of gyration, radiative losses

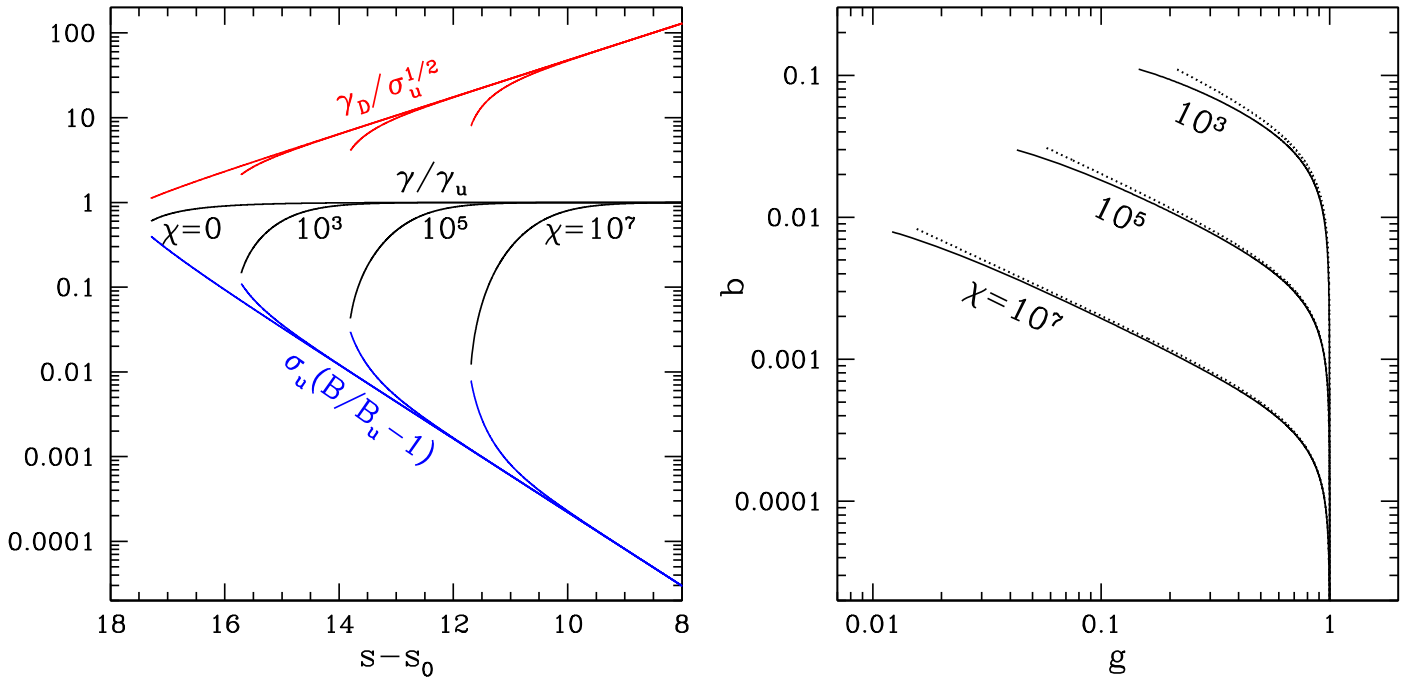

Figure 8. Shock structure solution in the approximation of steady $e^{\pm}$ streams (viewed in the shock rest frame). The solution was calculated for $\chi = 0$, 10^3, 10^5, and 10^7. Left: for each case, we show $g(s)$ (black), $b(s)$ (blue), and $\gamma_D(s) = B/\sqrt{B^2 - E^2}$ normalized to $\sqrt{\sigma_u}$ (red). Flows with $\chi \gg 1$ radiate most of their energy before developing gyration; this results in $\gamma \ll \gamma_u$. Each solution ends where the downstream begins, i.e., where the $e^{\pm}$ streams develop gyration (then they become unstable, thermalizing their gyration energy in the drift frame). The drift Lorentz factor γ_D at the end point represents the downstream speed relative to the shock. Right: relation between g and b. The approximate analytical solution (Equation (B19)) is shown by the dotted curve and compared with accurate numerical integration (solid curve). Both numerical and analytical solutions are plotted until the $e^{\pm}$ streams reach $|\tilde{\beta}_x| = 1/2$, which corresponds to 60° gyration in the drift frame.

become normal synchrotron emission, which carries away zero momentum in the drift frame and no longer affects the bulk motion of the plasma.

We set the end point s_{end} of the coherent stream solution where the $e^{\pm}$ velocity vector in the drift frame, $\tilde{\boldsymbol{\beta}}_{\pm}$, has rotated by 60° ($\tilde{\beta}_x^{\pm} = \mp 1/2$). We define the approximate downstream quantities at this point. In particular, the drift Lorentz factor $\gamma_D = B/\sqrt{B^2 - E^2}$ at s_{end} approximately represents the Lorentz factor of the downstream relative to the shock, γ_d. Figure 4 shows the measured γ_d as a function χ.[10] For shocks with $\chi \gg 1$, the numerical result approximately follows the relation $\gamma_d/\sqrt{\sigma_u} \approx \chi^{1/7}$. It can also be derived analytically, as shown below.

B.4. Approximate Analytical Solution

The radiation-reaction transition admits a simple analytical description. A key fact simplifying the analytical approximation is that the term $w\chi\mathcal{R}$ in dw/ds (Equation (B8)) remains small throughout the shock transition. When this term is neglected, the equations for dw/ds and db/ds can be integrated for $w(b)$. The resulting relation $w(b)$ is the same as in the adiabatic solution given in Equation (B12). At $b \gg \epsilon$, it gives

$$w \approx \frac{b^2}{2}. \tag{B15}$$

Note also that the functions $\mathcal{R}$ and $\mathcal{F}$ entering the dynamical equations can be simplified. Neglecting the small terms $\propto\epsilon$ in Equation (B11), we obtain

$$\mathcal{R} \approx (w + gb)^2 \approx b^2\left(\frac{b}{2} + g\right)^2, \qquad \mathcal{F} \approx \frac{b}{\sqrt{g}}, \tag{B16}$$

where we used $w/\sigma_u \ll g$ to further simplify $\mathcal{F}$. Like Equation (B15), Equations (B16) approximately hold at $b \gg \epsilon$, practically throughout the entire shock transition.

With these preliminary remarks, we can describe the shock structure as follows. The shock has two zones:

Adiabatic zone ($g_1 < g < 1$). Here, the upstream flow enters the shock by following the solution (B12) for $g(b)$ and $w(b)$. The adiabatic zone ends when the radiative term becomes important in dg/ds (Equation (B8)), i.e., when $\chi\mathcal{R}$ becomes comparable to $\mathcal{F}$.

[10] Different choices for the end point definition (up to $\tilde{\beta}_x^{\pm} = \mp\tilde{\beta} \approx 1$) affect the measured γ_d by tens of percent, a modest correction compared with the main effect—radiation reaction before the development of gyration. The exact γ_d may be found with the full kinetic plasma simulation of the shock.

This occurs at $b_1 \sim \chi^{-1} \ll 1$. By this moment, the flow has lost only a small fraction $1 - g_1$ of its initial kinetic energy, which went into the slight compression of the magnetic field,

$$b_1 \approx 1 - g_1 \approx \frac{1}{\chi} \ll 1. \tag{B17}$$

Radiation-reaction zone ($g_2 \ll g \ll g_1$). Here, $\chi\mathcal{R} \gg \mathcal{F}$, and the flow deceleration is controlled by radiative losses. The solution for $b(g)$ can be found from the ratio of equations $dg/ds \approx -\chi\mathcal{R}$ and $db/ds = w$ (Equation (B8)). Using $\mathcal{R} \approx g^2b^2$ (this holds because $gb \gg w$, as verified below), we find

$$\frac{3}{4}\chi(b^2 - b_1^2) \approx g^{-3/2} - g_1^{-3/2}. \tag{B18}$$

We have chosen the integration constant so that $b(g_1)$ matches b_1 (Equation (B17)). However, this boundary condition becomes unimportant in the radiation-reaction zone, where $b \gg b_1$ and $g^{-3/2} \gg 1 - g_1$. So, the solution simplifies to

$$\frac{3}{4}\chi b^2 \approx g^{-3/2} - 1. \tag{B19}$$

This result is in excellent agreement with the accurate numerical solution (Figure 8).

The key approximation of this derivation was the neglect of $w\chi\mathcal{R}$ compared with $b\mathcal{F}$ in dw/ds (Equation (B8)). Evaluating the ratio

$$\frac{w\chi\mathcal{R}}{b\mathcal{F}} \approx \frac{2}{3}\, g(1 - g^{3/2}), \tag{B20}$$

one can see that it has a maximum of $(2/5)^{5/3} \approx 0.2$, a small value. One can also verify that $w/\sigma_u \ll g$ and $gb/w \approx 2g/b \approx 3\chi\mathcal{R}/\mathcal{F} \gg 1$; this vindicates the approximations $\mathcal{F} \approx b/\sqrt{g}$ and $\mathcal{R} \approx g^2b^2$.

The radiation-reaction zone ends where the term $\chi\mathcal{R}$ stops being dominant in dg/ds. This transition may be defined, e.g., by $\chi\mathcal{R} \approx 2\mathcal{F}$, which occurs at

$$g_2 \approx \left(\frac{3}{\chi}\right)^{2/7} \ll 1. \tag{B21}$$

As the radiation-reaction zone ends at $g \sim g_2$, the gyrating downstream begins. The development of gyration can be seen by transforming the e^+ stream velocity to the drift frame. We find the drift speed $\beta_D = E/B = \beta_u B_u/B$ and then the drift Lorentz factor

$$\gamma_D = \frac{B}{\sqrt{B^2 - E^2}} = \frac{1 + b/\sigma_u}{\sqrt{(1 + b/\sigma_u)^2 - \beta_u^2}} \approx \sqrt{\frac{\sigma_u}{2b}}. \tag{B22}$$

The particle Lorentz factor measured in the drift frame is

$$\tilde{\gamma} \approx \gamma_D(\gamma - \beta_D u_z) \approx \gamma_u \sqrt{\frac{b}{2\sigma_u}}\left(g + \frac{b}{2}\right). \tag{B23}$$

The $e^\pm$ streams are ultrarelativistic in the drift frame; i.e., $\tilde{\beta} \approx 1$, and $\tilde{\beta}_x \approx \sin\tilde{\psi}$ determines their deflection angles $\pm\tilde{\psi}$ from the z-axis. Taking into account that $\tilde{u}_x = u_x$, we find

$$\tilde{\beta}_x = \frac{u_x}{\tilde{\gamma}} = -\frac{\gamma\mathcal{F}}{\tilde{\gamma}\sqrt{\sigma_u}} \approx -\frac{\sqrt{2gb}}{g + b/2}. \tag{B24}$$

Substituting here the approximate solution (B19) for $b(g)$, one can see the development of gyration with decreasing g. In particular, one can see that $|\tilde{\beta}_x| \sim 1$ at $g \sim g_2$, confirming the formation of a gyrating downstream.

Finally, we can evaluate the shock jump condition of main interest: the downstream Lorentz factor in the shock rest frame, γ_d (denoted as $\gamma_d' = \gamma_{sh|d}$ in the main text). It approximately equals $\gamma_D(g_2)$, which gives

$$\frac{\gamma_d}{\sqrt{\sigma_u}} \approx \chi^{1/7} \qquad (\chi \gg 1). \tag{B25}$$

ORCID iDs

Andrei M. Beloborodov https://orcid.org/0000-0001-5660-3175

References

Alsop, D., & Arons, J. 1988, PhFl, 31, 839
Arons, J., & Barnard, J. J. 1986, ApJ, 302, 120
Beloborodov, A. M. 2013, ApJ, 777, 114
Beloborodov, A. M. 2017, ApJL, 843, L26
Beloborodov, A. M. 2020, ApJ, 896, 142
Beloborodov, A. M. 2021a, ApJL, 922, L7
Beloborodov, A. M. 2021b, ApJ, 921, 92
Beloborodov, A. M. 2022, PhRvL, 128, 255003
Beloborodov, A. M. 2023, arXiv:2307.12182
Blaes, O., Blandford, R., Goldreich, P., & Madau, P. 1989, ApJ, 343, 839
Bransgrove, A., Beloborodov, A. M., & Levin, Y. 2020, ApJ, 897, 173
Chen, A. Y., Yuan, Y., Li, X., & Mahlmann, J. F. 2022, arXiv:2210.13506
Erber, T. 1966, RvMP, 38, 626
Falcke, H., & Rezzolla, L. 2014, A&A, 562, A137
Heyl, J. S., & Hernquist, L. 2005, MNRAS, 362, 777
Kaspi, V. M., & Beloborodov, A. M. 2017, ARA&A, 55, 261
Kennel, C. F., & Pellat, R. 1976, JPlPh, 15, 335
Landau, L. D., & Lifshitz, E. M. 1975, The Classical Theory of Fields (Oxford: Pergamon)
Langdon, A. B., Arons, J., & Max, C. E. 1988, PhRvL, 61, 779
Lehner, L., Palenzuela, C., Liebling, S. L., Thompson, C., & Hanna, C. 2012, PhRvD, 86, 104035
Levinson, A. 2022, MNRAS, 517, 569
Li, X., Beloborodov, A. M., & Sironi, L. 2021, ApJ, 915, 101
Li, X., Zrake, J., & Beloborodov, A. M. 2019, ApJ, 881, 13
Longair, M. S. 1994, High Energy Astrophysics, Vol. 2 (Cambridge: Cambridge Univ. Press)
Lyubarsky, Y. 2014, MNRAS, 442, L9
Lyubarsky, Y. 2021, Univ, 7, 56
Lyubarsky, Y. E. 2003, MNRAS, 339, 765
Mahlmann, J. F., Philippov, A. A., Mewes, V., et al. 2023, ApJL, 947, L34
Most, E. R., Nathanail, A., & Rezzolla, L. 2018, ApJ, 864, 117
Most, E. R., & Philippov, A. A. 2020, ApJL, 893, L6
Nättilä, J., & Beloborodov, A. M. 2022, PhRvL, 128, 075101
Parfrey, K., Beloborodov, A. M., & Hui, L. 2013, ApJ, 774, 92
Philippov, A., & Kramer, M. 2022, ARA&A, 60, 495
Sironi, L., Plotnikov, I., Nättilä, J., & Beloborodov, A. M. 2021, PhRvL, 127, 035101
Thompson, C. 2023, MNRAS, 519, 497
Thompson, C., & Blaes, O. 1998, PhRvD, 57, 3219
Thompson, C., & Duncan, R. C. 1996, ApJ, 473, 322
Troischt, P., & Thompson, C. 2004, PhRvD, 70, 124030
Yuan, Y., Beloborodov, A. M., Chen, A. Y., & Levin, Y. 2020, ApJL, 900, L21
Yuan, Y., Levin, Y., Bransgrove, A., & Philippov, A. 2021, ApJ, 908, 176