Sub-photospheric Shocks in Relativistic Explosions

Andrei M. Beloborodov
Physics Department and Columbia Astrophysics Laboratory, Columbia University, 538 West 120th Street, New York, NY 10027, USA
Received 2016 April 10; revised 2017 January 1; accepted 2017 January 11; published 2017 April 3

Abstract
This paper examines the mechanism of internal shocks in opaque relativistic outflows, in particular in cosmological gamma-ray bursts. The shocks produce neutrino emission and affect the observed photospheric radiation from the explosion. They develop from internal compressive waves and can be of different types depending on the composition of the outflow. (1) Shocks in “photon gas,” with negligible plasma inertia, have a unique structure determined by the force-free condition—zero radiation flux in the plasma rest frame. Radiation dominance over plasma inertia suppresses the formation of collisionless shocks mediated by collective electromagnetic fields. (2) If the outflow is sufficiently magnetized, a strong collisionless subshock develops, which is embedded in a thicker radiation-mediated structure. (3) Waves in outflows with a free neutron component lead to dissipation through nuclear collisions. At large optical depths, shocks have a thickness comparable to the neutron free path, with embedded radiation-mediated and collisionless subshocks. The paper also presents first-principles simulations of magnetized flows filled with photons, demonstrating the formation of shocks and their structure. Simple estimates show that magnetized sub-photospheric shocks are efficient producers of photons and have a great impact on the observed photospheric radiation. The shock structure changes as the outflow expands toward its photosphere. The dissipation is accompanied by strong $e^\pm$ pair creation, and the $e^\pm$-dressed shock carries the photosphere with it up to two decades in radius, emitting a strong pulse of nonthermal radiation.

Key words: gamma-ray burst: general – magnetohydrodynamics (MHD) – neutrinos – radiation mechanisms: non-thermal – radiative transfer – shock waves

1. Introduction
Astrophysical explosions and jets generate shock waves, which produce radiation. Their radiative properties are determined by the dissipation mechanism that sustains the velocity jump in the shock and by its ability to generate nonthermal particles. This paper examines the mechanism of internal shocks in gamma-ray bursts (GRBs) that occur before the GRB jets become transparent to radiation. The approach and some of the results may also be of interest for other explosions, e.g., in novae or supernovae.

1.1. Internal Shocks in GRB Jets
The main features of GRB explosions may be summarized as follows: the outflow is relativistic, it carries magnetic fields frozen in fully ionized plasma, and a large fraction of its energy is carried by neutral particles—photons and free neutrons. GRB outflows start very opaque near the central engine of the explosion and become transparent at a large “photospheric” radius $R_p$. Internal shocks can develop below and above the photosphere.

Early works proposing internal shocks in GRBs focused on shocks above the photosphere (Rees & Mészáros 1994; Kobayashi et al. 1997; Daigne & Mochkovitch 1998). They can only be collisionless, i.e., mediated by collective electromagnetic fields. Their mechanism has been studied in detail using particle-in-cell simulations, and it was found that the presence of transverse magnetic fields renders the shock unable to accelerate particles (Sironi & Spitkovsky 2011): the postshock electron–ion plasma is in a two-temperature state, $T_e < T_i$, with electrons and ions forming nearly Maxwellian distributions. This may change, however, if the electron–ion outflow is loaded with $e^\pm$ plasma. Ultrarelativistic shocks in $e^\pm$-loaded plasma with transverse magnetic field were found to be capable of accelerating positrons (Hoshino et al. 1992; Amato & Arons 2006; Stockem et al. 2012).

Shocks in opaque plasma below the photosphere $R_p$ are less explored and may be key to understanding GRB emission (Mészáros & Rees 2000b; Pe’er et al. 2006; Beloborodov 2010; Levinson 2012). There is significant evidence that GRB radiation is mainly produced below the photosphere (Ryde et al. 2011; Beloborodov 2013; Yu et al. 2015), and detailed simulations of radiative transfer in opaque heated jets give spectra consistent with GRB observations (Vurm & Beloborodov 2016). Internal shocks provide a plausible mechanism for sub-photospheric heating invoked by these models.

Internal shocks may result from the fast variability of the central engine or the outflow interaction with the progenitor star (e.g., Lazzati et al. 2013; Ito et al. 2015). At later stages of ballistic expansion with a high Lorentz factor $\Gamma$, internal shocks at radius $r$ can develop from velocity variations on a scale $L \sim r/\Gamma$ (measured in the outflow rest frame). This scale also sets the characteristic optical depth seen by photons in the expanding outflow, $\tau_T \sim \sigma_T n_e L$, where $\sigma_T$ is the Thomson cross section and $n_e$ is the proper density of electrons and positrons.

The present paper is motivated by the following questions.

1) Can sub-photospheric shocks be collisionless? This is assumed in models of TeV neutrino emission from the jet–progenitor interaction (Razzaque et al. 2003), but the assumption is questionable (Murase & Ioka 2013).

2) Is the shock capable of producing high-energy particles? The presence of high-energy electrons at large optical depths would have a strong effect on the photospheric radiation (Pe’er & Waxman 2004; Beloborodov 2010; Vurm & Beloborodov 2016).

3) How does the shock evolve as it emerges from the photosphere and what is its observational appearance?
1.2. Radiation-mediated Shocks (RMSs)

Since GRB jets carry a large number of photons per electron, sub-photospheric shocks are naturally expected to be mediated by radiation (Levinson 2012). Then dissipation and the profile of the velocity jump are controlled by photon scattering.

Basic features of RMSs were studied in the 1950s (see Zeldovich & Raizer 1966 and references therein). The RMS propagation is sustained by radiation diffusion: radiation generated by the shock diffuses upstream and preheats the upstream gas. This creates a pressure gradient, a kind of a "pillow" that allows the gas to smoothly decelerate, avoiding the collapse of the shock thickness to the collisionless scale (the ion Larmor radius).

The first RMS models assumed that radiation is everywhere in local thermodynamic equilibrium. This assumption can be strongly violated in astrophysical explosions, because the timescale to establish thermodynamic equilibrium can be much longer than the time it takes the gas to cross the shock. Models relaxing this assumption of equilibrium have been developed and applied to supernova shock breakout (e.g., Weaver 1976; Sapir et al. 2013). The RMS model was also extended to relativistic shocks (Levinson & Bromberg 2008; Budnik et al. 2010; Bromberg et al. 2011; Levinson 2012). The highest temperature achieved in an RMS depends on the number carried by the upstream through the shock. Levinson (2012) emphasized the low efficiency of photon production by the RMS in GRB jets and developed a shock model with a conserved photon number.

The RMS thickness is large, comparable to or larger than the photon mean free path. This inhibits diffusive acceleration of charged particles. In particular, electrons radiate energy faster than they can cross the shock. The RMS is only capable of a slow diffusive acceleration of photons up to the MeV band (in the shock frame).

The RMS picture of internal shocks, however, has a few caveats. Previous work did not take into account that the outflow is magnetized, and the magnetic field can change the RMS structure and the dissipation mechanism. In addition, GRB explosions are expected to carry free neutrons; their collisions can play a key role in shaping the shock waves at large optical depths and offer an additional mechanism for producing high-energy particles and neutrinos.

1.3. Outline of the Paper

We begin with the basics of shock formation. Section 2 examines how a supersonic compressive wave steepens and launches a pair of shock waves. We first describe shock formation in a polytropic gas using the hydrodynamic approximation (zero mean free path of all particles and photons). Then we relax this approximation and discuss the role of photon diffusion in the formation of RMSs and collisionless shocks. We consider a "cold" gas with sound speed $c_0 \ll c$ and formulate two conditions for the immediate RMS formation (versus formation of a collisionless shock). We also discuss flows with large $c_0$, including the extreme regimes where the flow inertia is dominated by radiation ($c_0 = c/\sqrt{3}$) or magnetic fields ($c_0 = c$).

Section 3 describes the general jump conditions for shocks in media with any thermal pressure and magnetization. A moderate magnetization of the flow changes the jump conditions and we argue that this leads to the formation of a thin collisionless subshock, even at large optical depths. We evaluate the region in the parameter space where a strong collisionless subshock must exist.

Section 4 focuses on shocks in “photon gas” with subdominant magnetic fields and negligible plasma inertia ($c_0 = c/\sqrt{3}$). This regime may occur in GRB explosions in their early stages, during the jet breakout and its acceleration by radiation pressure. We use a self-consistent simulation of time-dependent radiative transfer and obtain the solution for the shock structure.

Then Section 5 presents the RMS structure at later stages when the plasma inertia becomes important. As the main tool, we use direct Monte-Carlo simulations of time-dependent radiative transfer coupled with the flow dynamics. We first investigate RMS formation in a weakly magnetized flow, and discuss the effect of bulk Comptonization and $e^\pm$ creation inside the shock front. Then we turn to shocks in a magnetized fluid and demonstrate the formation of a strong collisionless subshock embedded in an RMS, as anticipated in Section 3.

Section 6 investigates how plasma heating in a collisionless (sub)shock results in “breeding” of $e^\pm$ pairs. Section 7 discusses the emergence of an $e^\pm$-dressed shock from the photosphere.

Section 8 describes shocks in outflows with a free neutron component. Due to their large free paths, neutrons introduce a large effective viscosity. Nuclear collisions produce ultrarelativistic $e^\pm$ pairs and neutrinos. They sustain broad shock fronts until the jet reaches the neutron decoupling radius $R_n$, where most neutrons begin to flow freely without collisions.

The results and their implications for GRB models are discussed in Section 9.

2. Formation of Shocks

2.1. Ballistic Approximation and Caustics

Consider an outflow with internal supersonic motions. Such motions can be approximately described as ballistic: each fluid element is moving with a constant velocity. It is well known that ballistic flows create caustics—surfaces where density diverges (in cosmology such surfaces are called Zeldovich pancakes).

The flow near the caustic is approximately plane-parallel (one-dimensional). It is convenient to view the flow in the rest frame of the caustic and choose the $x$-axis normal to it, so that the flow converges toward $x = 0$ along the $x$-axis with velocity $v(t, x)$. Since $\beta = v/c$ may approach unity in relativistic flows, it is useful to introduce dimensionless momentum $p = \gamma \beta$ where $\gamma = (1 - \beta^2)^{-2}$. Velocity is related to $p$ by

$$v = \frac{cp}{(1 + p^2)^{1/2}}.$$ (1)

A simple example of a converging flow is provided by an initial arctan profile,

$$p_0(x) = -p_{max} \frac{2}{\pi} \arctan \frac{x}{L}.$$ (2)

The ratio $p_{max}/L$ describes the initial steepness of the wave, and $p_{max}$ describes its amplitude. The wave is nonrelativistic if $p_{max} \ll 1$. The characteristic timescale of the profile evolution is $t_c = L/c_{\mathsf{p max}}$. On this timescale the ballistic wave steepens (Figure 1) and the caustic forms at $x = 0$, where $-\partial v/\partial x$ is maximum.
The density of the ballistic flow diverges at the caustic. Its evolution is determined by the relation

\[ x(x_0, t) = x_0 + v_0 t, \]

where \( x_0 \) is the initial position of the fluid slab \( dx_0 \) at times \( t_0 \ll t_c \) and \( v_0(x) = v_0(x) \) is the initial velocity profile. The slab \( dx_0 \) is contracting by the factor \( \frac{dx_0}{dx_0} = 1 + (dv_0/dx_0)t \).

Therefore, the evolution of baryon density \( \rho \) is described by

\[ \rho(x_0, t) = \rho_0(x_0) \left( 1 + v_0 t \right), \quad v_0 = \frac{dv_0}{dx_0}, \]

where \( \rho_0(x_0) = \rho_0(x) \) is the density at \( t_0 \ll t_c \). The compression rate is highest at \( x = 0 \) and here density diverges at time

\[ t_c = \left( \frac{dv_0}{dx_0} \right)^{-1} \biggr|_{x_0=0} = \frac{\pi L}{2c p_{\text{max}}}. \]

At this moment, \( v(x) \) becomes discontinuous at \( x = 0 \).

### 2.2. Pressure Buildup in the Converging Flow

True caustics form in flows with zero pressure. A small initial pressure \( P_0 \approx 0 \) qualitatively changes the picture: it can be strongly amplified in the converging flow near \( x = 0 \) and the generated pressure gradient stops the flow.

The deceleration of the converging flow around \( x = 0 \) accelerates the steepening of the velocity profile on each side of the caustic (Figure 2). As a result, at some time \( t_* \) and locations \( \pm x_* \) two shocks form and continue to propagate away from \( x = 0 \). The type of the nascent shocks depend on the physical conditions in the region \( x \sim x_* \). Below we discuss the pressure buildup in the converging flow, then estimate the location of shock formation \( x_* \) and the corresponding maximum compression.

One source of pressure is the thermal motion of plasma particles. It grows in the converging flow, but its contribution to the total pressure is limited by fast radiative cooling, which converts plasma heat to radiation. In local thermodynamic equilibrium, radiation strongly dominates the heat capacity of GRB jets, because the photon density \( n_\gamma \) greatly exceeds the plasma density. At small radii, where the \( e^\pm \) population is in annihilation equilibrium with Planck radiation, one finds \( n_+ / n_\gamma \approx (KT/m_e c^2)^{-3/2} \exp(-m_e c^2 / KT) \) (Svensson 1984); the \( e^\pm \) abundance is decreasing exponentially in the expanding and adiabatically cooling jet. At larger radii, the particle density (ions, \( e^- \), or \( e^+ \)) does not exceed \( \sim 10^{-3}n_\gamma \), and, in local thermodynamic equilibrium, this implies a plasma pressure \( P_{\text{pl}} \ll P_{\text{rad}} \). Here we examine compressive waves in a medium that is initially not too far from thermal equilibrium\(^1\) and thus has \( P_{\text{pl}} \ll P_{\text{rad}} \). Then the two main sources of pressure that can be amplified in the wave are radiation and the transverse magnetic field.

Magnetic fields are expected to carry a significant fraction \( \epsilon_B \) of the jet energy. Comparison of theoretical GRB spectra with observations suggests \( \epsilon_B \sim 0.01-0.1 \) (Vurm & Beloborodov 2016). The jet plasma is an excellent conductor, so the magnetic field is frozen in it and advected by the flow. In an internal compressive wave, the frozen transverse field is compressed together with the plasma: \( B \propto \rho \) or \( B \propto \rho \), where \( B = B / \gamma \) is the magnetic field measured in the fluid frame, \( \rho = \rho / \gamma \) is the proper density of the fluid, and \( \gamma \) is its Lorentz

\(^1\) Shocks create strong deviations from thermodynamic equilibrium, and these deviations become long-lived in the region of moderate optical depth, around and above the photosphere. New shocks in this region will develop in the plasma with hot (thermally decoupled) ions, preheated by previous shocks.
factor. The magnetic pressure grows in the converging flow as\footnote{\(P_B\) and \(\bar{\rho}\) are measured in the same (fluid) frame. Pressure and internal energy density are always measured in the fluid frame and we omit the tilde to simplify notation.}:
\[
P_B = \frac{B^2}{8\pi} \propto \bar{\rho}^2. \tag{6}
\]

The growth of radiation pressure \(P_{\text{rad}}\) depends on its ability to diffuse out of the compressed region, which depends on the optical depth. If the flow is sufficiently opaque to photons, the radiation will be trapped and
\[
P_{\text{rad}} = \frac{U_{\text{rad}}}{3} \propto \bar{\rho}^{4/3} \quad \text{(trapped radiation).} \tag{7}
\]
In the opposite limit, when the compressed region is transparent to photons, there is no significant amplification of \(P_{\text{rad}}\).

Equations (6) and (7) both have a polytropic form \(P \propto \bar{\rho}^\alpha\), with \(\alpha = 2\) or 4/3. A similar relation could also be used for compressive waves in a medium that is far from thermal equilibrium with radiation and filled with hot, thermally decoupled, ions (protons). The proton pressure in the compressive wave follows the relation \(P_p \propto \bar{\rho}^\alpha\) with \(\alpha = 5/3\) as long as the proton temperature is nonrelativistic, \(kT_p \lesssim m_p c^2\).

2.3. Shock Formation in Nonrelativistic Polytropic Gas

Let us first consider a nonrelativistic gas, \(P \ll \bar{\rho} c^2\). Suppose that initially the gas has uniform pressure \(P_0\) and density \(\bar{\rho}_0\), and is set in motion with velocity \(v_0(x)\) that corresponds to momentum profile \(p_0(x)\) given, e.g., by Equation (2). We assume that the peak of velocity profile \(v_{\text{max}} = c p_{\text{max}} (1 + p_{\text{max}}^2)^{-1/2}\) is much greater than the sound speed \(c_0 = (\alpha P_0/\bar{\rho}_0)^{1/2}\). In the ballistic approximation, the profile would develop a caustic at \(x = 0\) at time \(t\). We wish to know how the finite pressure changes the flow dynamics, in particular what is the maximum pressure achieved in the compressed region before a shock forms, and where the shock formation occurs.

Even if the compressive wave is relativistic, \(p_{\text{max}} \gtrsim 1\), the condition \(c_0 \ll c\) implies that the shocks form not far from \(x = 0\) where \(v_0(x)\) is nonrelativistic. Therefore, the shock formation can be examined using Newtonian hydrodynamics around \(x = 0\), so we will use \(\gamma = 1\), \(\bar{\rho} \approx \rho\), and \(\mathcal{B} \approx B\).

The evolution of the gas is convenient to view on the \(x-t\) plane (Figure 3). Each streamline is described by \(x(x_0, t)\) where \(x_0\) is the Lagrangian coordinate—the position at \(t = 0\). Initially, a small fraction of gas is in the subsonic region \(|v_0| < c_0\) near \(x = 0\). The streamlines that start outside this region are initially supersonic and eventually become subsonic.

There is a critical Lagrangian coordinate \(x_0^*\). Streamlines that start at \(|x_0| < x_0^*\) will become subsonic without a shock: the compressed gas is gradually decelerated as its specific kinetic energy \(v_0^2/2\) gets transformed into enthalpy \((U + P)/\rho = c_s^2/(\alpha - 1)\), where \(c_s^2 = c_0^2 (\rho/\bar{\rho}_0)^{\alpha - 1}\) is the local speed of sound. This “compressive deceleration” to a subsonic speed occurs when the compression factor \(\rho/\bar{\rho}_0\) satisfies
\[
\frac{c_0^2}{\alpha - 1} \left(\frac{\rho}{\bar{\rho}_0}\right)^{\alpha - 1} \approx \frac{v_0^2}{2}. \tag{8}
\]

Approximating the streamline before this moment as ballistic, one can estimate \(\rho/\bar{\rho}_0 = 1 + v_0^2(x_0)/v_{\text{dec}}^2\). Therefore, the deceleration time \(t_{\text{dec}}(x_0)\) at which the streamline with Lagrangian coordinate \(x_0\) becomes subsonic may be estimated from the condition
\[
(1 + v_0^2 t_{\text{dec}})^{\alpha - 1} \approx \frac{2c_0^2}{(\alpha - 1) v_0^2}. \tag{9}
\]

The corresponding location on the streamline is
\[
x_{\text{dec}} \approx x_0 + v_0 t_{\text{dec}}. \tag{10}\]

The smooth compressive deceleration is possible only for streamlines with sufficiently small \(|x_0| < x_0^*\). For large \(|x_0|\) one finds \(t_{\text{dec}} > |x_0|/v_0^2\), and the compressive deceleration becomes impossible—the ballistic flow does not have a chance to compress enough before it hits the existing subsonic region near \(x = 0\). Then the deceleration occurs through a shock.

The critical Lagrangian coordinate \(x_0^*\) at which the shock forms is given by (see the Appendix)
\[
\frac{x_0^*}{L} \approx 3 \left(\frac{c_0}{c_{\text{max}}}\right)^{1/\alpha}. \tag{11}\]

The compression of gas with the Lagrangian coordinate \(x_0^*\) is determined by Equation (8) with \(v_0\) evaluated at \(x_0^*\). In the limit of \(x_0^* \ll L\), the compression along this streamline is given by
\[
\frac{\rho}{\bar{\rho}} = \frac{x_0^*}{x_*} \approx \left(\frac{c_{\text{max}}}{c_0}\right)^{2/\alpha}. \tag{12}\]

Figure 3. The streamlines on the spacetime diagram. The gas has the initial velocity profile \(v_0(x_0)\) given by Equation (2) with \(p_{\text{max}} = 0.7\), the initial sound speed \(c_0 = 0.1 c\), and the adiabatic index \(\alpha = 4/3\). The boundary of the subsonic region \(v < c_s\) is shown by the red curves. The red curve is dashed where the deceleration to the subsonic speed occurs smoothly and solid where the deceleration occurs through a shock. The shock forms at \(t_s \approx 2.9 L/c\) and \(x_s \approx 0.07 L\).
This determines the maximum pressure developed in the flow before the shock is launched,
\[
P_s = P_0 \left( \frac{\rho_s}{\rho_0} \right)^2 \sim P_0 \left( \frac{c_{\text{lim}}}{c_0} \right)^2 \sim P_{\text{max}}^2 \rho_0 c^2.
\]  
(13)

The maximum pressure is comparable to the peak kinetic energy density of the wave, even though \( P_s \) develops only in a small region near the caustic \( x \approx 0 \) where the flow momentum is much smaller than \( P_{\text{max}} \). This is because \( P_s \) is controlled by the curvature of the velocity profile (described by \( v_0^4(0) \), see the Appendix), which depends on \( P_{\text{max}} \).

At \( t_s \) and \( x_s \), the sound speed of the ballistic flow is not much below its bulk speed \( v \sim v_0(x_0^4) \), so the nascent shock is not strong. Then the shock propagates through the ballistic gas with increasing Lagrangian coordinate \( |x_0| \) where the upstream velocity \( v_0(x_0^4) \) is higher, and the shock compression ratio quickly approaches the strong-shock limit \((\alpha + 1)/(\alpha - 1)\).

2.4. Shock Formation in Relativistic Polytropic Gas

Shock formation in relativistic gas with \( P \gg \rho c^2 \) may be examined in a similar way. This regime occurs in relativistic explosions at small radii where radiation dominates the gas inertia. Then \( c_s^2 = c^2/3 \) (Landau & Lifshitz 1959), and one must consider relativistic compressive waves with \( v_{\text{max}} > c_0 \).

The flow is initially subsonic in the zone where \( |p_0| < 2^{-1/2} \). Outside this zone the flow is approximately ballistic and its density grows with time as \((1 + v_0^4/t)^{-1}\), where \( v_0 = c(d\rho_0/dx_0)^{1/2} \). The compressive deceleration of the relativistic gas is quite efficient: a large fraction of the bulk kinetic energy is converted to enthalpy when the gas is compressed by only a factor of \( \sim 2 \).

However, even such a moderate compression is difficult to achieve in the relativistic ballistic flow, because the gas with \( \gamma_0 \gg 1 \) (\( v_0 \approx c \)) has a small \( v_0^4 \) and hence it is compressed slowly. The maximum time allowed for ballistic compression is \( x_0/c \) and the corresponding maximum compression factor is
\[
(1 + v_0^4/x_0^4/c)^{-1} \approx 1 \quad \text{if} \quad |p_0| \gg 1.
\]  
(14)

Gas with a relativistic \( p_0 \) ballistically hits the subsonic region before it has a chance for compressive deceleration. Thus, the shock must form at Lagrangian coordinate \( x_0^* \) such that \( |p_0(x_0^*)| \sim 1 \), not far from the boundary of the initial subsonic zone \( |p_0| = 2^{-1/2} \). The time and location of shock formation are \( t_s \sim x_0^*/c \) and \( x_\text{shock} \sim x_0^*/2 \). The shock forms with a mildly relativistic amplitude; it becomes ultrarelativistic when it propagates into the gas converging with \( |p_0| \gg 1 \).

One can also consider shock formation in a magnetically dominated gas with \( P_B \gg \rho c^2 \) and \( P_B \gg \rho p_0 \). Then \( c_0 \approx c \) and it is convenient to define \( \gamma_0 = 1 - c_0^2/c^2 = \gamma_{c0}^{-1/2} \). Shocks form in compressive waves with \( v_{\text{max}} > c_0 \), which corresponds to a Lorentz factor \( \gamma_{\text{max}} > \gamma_{c0}^{-1} \).

In the limit of strong magnetization, the sound speed becomes equal to \( c \) and shocks do not form. In this case, the dynamic equations read \( \partial_\mu T^{\mu \nu} = 0 \) with the stress–energy tensor components
\[
T^\eta = \frac{B^2 + E^2}{8\pi}, \quad T^{\eta \xi} = \frac{EB}{4\pi}, \quad T^{\eta \xi} = \frac{B^2 + E^2}{8\pi}
\]  
(15)

(the magnetic field \( B \) is assumed to lie in the \( y-z \) plane perpendicular to the fluid velocity). The neglect of the plasma contribution to \( T^{\mu \nu} \) defines so-called force-free electrodynamics, where plasma serves only to conduct electric currents demanded by \( \nabla \times B \) and supplies no inertia. The plasma velocity \( \nu = \beta c \) is related to the electric field by \( E + \nu \times B/c = 0 \) and \( \beta = E/B \). Adding and subtracting the energy and momentum conservation laws,
\[
\frac{\partial T^{\mu \nu}}{\partial t} + c \frac{\partial T^{\mu \nu}}{\partial x^\mu} = 0, \quad \frac{\partial T^{\mu \nu}}{\partial t} + c \frac{\partial T^{\mu \nu}}{\partial x^\nu} = 0,
\]  
(16)

one obtains
\[
\frac{\partial u_\pm}{\partial t} \pm c \frac{\partial u_\pm}{\partial x} = 0, \quad u_\pm = (1 \pm \beta)B.
\]  
(17)

The initial profiles of \( u_\pm(0, x) = f_\pm(x) \) determine \( u_\pm(t, x) = f_\pm(x + ct) \). This gives explicit solutions for \( B = (u_+ + u_-)/2 \) and \( \beta = (u_+ - u_-)/(u_+ + u_-) \), demonstrating their smooth behavior, with no caustics or shocks.

2.5. Radiation Diffusion and Formation of an RMS

Radiation diffusion is an essential ingredient of an RMS, since it is the mechanism of shock propagation. However, diffusion that is too fast would let radiation escape, inhibiting the RMS formation. A shock wave is usually assumed to be radiation-mediated if two conditions are satisfied:

(A) The jump conditions give \( P \approx P_{\text{rad}} \) the downstream, so that a large fraction of energy generated by the shock is carried by radiation (Zeldovich & Raizer 1966).

(B) The medium has optical depth \( \tau > c/v_0 \), so that the shock generates radiation faster than it can diffuse out of the system of size \( L \). For instance, in a supernova explosion one could take \( L \) as the radius of the expanding ejecta (e.g., Tolstov et al. 2013).

In fact, these conditions do not guarantee that the shock is mediated by radiation. The velocity profile connecting the upstream and downstream may contain a “subshock”—a sharp jump mediated by the plasma on a scale much smaller than the photon free path to scattering. In nonrelativistic shocks \( v_0 \ll c \) satisfying conditions (A) and (B) the velocity profile is smooth, with no subshock (Zeldovich & Raizer 1966). However, in the relativistic case, \( v_0 \approx c \), a weak subshock was reported (Budnik et al. 2010). In addition, in the above condition (B) one should be careful with what is meant by the “size of the system.”

Consider a compressive wave of a mildly relativistic amplitude \( v_{\text{max}} \sim c \) and length \( \sim L \). A characteristic optical depth may be defined as
\[
\tau_L = L/\rho_0 \kappa,
\]  
(18)

where \( \kappa \) is the opacity of the gas. Suppose the unperturbed gas has a nonrelativistic sound speed \( c_0 \ll c \). Section 2.3 described how at time \( t_s \sim L/c \) two shocks form near the caustic, at the Lagrangian coordinate \( x_0^*/L \sim (c_0/c)^{1/\gamma} \). Thus, the region of shock formation has the optical depth
\[
\tau_s \sim \tau_L \left( \frac{c_0}{c} \right)^{1/\gamma}.
\]  
(19)
For a radiation-dominated flow, \( P \approx P_{\text{rad}} \) and \( c = 4/3 \). In this case, however, Equation (19) can only be used if radiation is trapped, i.e., unable to diffuse out of the region of shock formation on the timescale \( x_s/v_0^2 \). This requires
\[
\tau_* \gg \frac{c}{v_0^2}, \tag{20}
\]
where \( v_0^2 \approx c/\tau_{\text{L}} \) is the flow velocity upstream of the nascent shock. The trapping condition is satisfied if
\[
\frac{c_0}{c} \gg \tau_{\text{L}}^{-2/3}. \tag{21}
\]
For the flow with the upstream pressure \( P_0 \) dominated by radiation, one can use the relation
\[
\frac{c_f}{c} \approx \left( \frac{w}{3} \right)^{1/2}, \quad w = \frac{U_{\text{rad}} + P_{\text{rad}}}{\rho c^2}. \tag{22}
\]
Then condition (21) may also be written as \( w \gg 3\tau_{\text{L}}^{-4/3} \). If this condition is satisfied, a propagating jump in radiation pressure will develop at \( t_* \), and the nascent shock will be mediated by photons, i.e., an RMS will be launched.

The RMS velocity profile is shaped by the competition between advection of radiation through the shock and its diffusion in the opposite, upstream direction. Therefore, the optical depth of the velocity jump \( \Delta \tau \) is regulated to
\[
\Delta \tau \sim \frac{c}{v_0}. \tag{23}
\]
The RMS propagation involves continual amplification of radiation advected through the shock—the result of photon scattering in the region of a steep velocity gradient. As the hot downstream photons diffuse back into the upstream, they experience “bulk Comptonization”—they are boosted in energy by a factor of \( \sim \sqrt{1 - v_f^2/c^2} \). As a result, the energy of radiation advected through the shock is amplified, as required by the jump conditions for a propagating shock.

Launching an RMS at \( t_* \) requires an initial buildup of radiation density \( \sim \rho v_f^2/2 \) near the shock front, which is possible only if the trapping condition (21) is satisfied. Otherwise, radiation leaks out of the compressed region to large distances \( x \). This may be viewed as a violation of RMS condition (B), because the effective “size of the system” during the shock formation is comparable to \( x_s \ll L \). Then the radiation pressure gradient is too weak to control the velocity profile of the flow. Radiation is unable to resist the steepening of the velocity profile, and the width of the velocity jump is quickly reduced to the ion Larmor radius, forming a collisionless shock mediated by the collective electromagnetic field. It may later evolve into an RMS, when the postshock region has accumulated a sufficient optical depth, if there is enough time for that in the expanding outflow, i.e., if the shock forms sufficiently far below the photosphere.

### 2.6. Critical Magnetization

When both magnetic field and radiation contribute to pressure, there are two contributions to the sound speed, \( c_s^2 = c_{\text{rad}}^2 + c_B^2 \). It is convenient to define the dimensionless enthalpy of the flow,
\[
w = \frac{P_{\text{rad}} + U_{\text{rad}}}{\rho c^2} = \frac{4P_{\text{rad}}}{\rho c^2}. \tag{24}
\]
A similar quantity for the magnetic field \( B \) in the fluid frame is
\[
\sigma = \frac{P_{\text{B}} + U_{\text{B}}}{\rho c^2} = \frac{B^2}{4\pi \rho c^2}. \tag{25}
\]
Suppose the optical depth is large so that the radiation trapping condition is satisfied. The type of the nascent shock is determined by whether the magnetic field or radiation dominates the pressure in the compressed region near the caustic, \( P_* = P_{\text{rad}}^* + P_{\text{B}}^* \). An RMS forms if \( P_* \) is dominated by radiation; otherwise, a collisionless shock is launched. Since \( P_* \) must be the same in either case (see Equation (13)), it is sufficient to compare the compressions needed to reach \( P_{\text{B}}^* \) with only magnetic or only radiation pressure: \( (P_{\text{rad}}/P_{\text{rad}})^{3/4} \) and \( (P_{\text{B}}/P_{\text{B}})^{1/2} \). This comparison gives an approximate condition for launching a collisionless shock in a cold \( (w_0 \ll 1) \) and opaque medium,
\[
\sigma_0 > \frac{w_0^{3/2}}{4\rho_{\text{max}}^{1/2}}. \tag{26}
\]
Note that this condition applies only to the nascent shock, at the point of maximum ballistic compression in the converging wave near the caustic. As the shock becomes stronger and continues to propagate into the ballistic flow where the upstream plasma is less compressed (and hence less magnetized), its type may change.

An established steady shock structure is determined by the parameters of its upstream. The first step in the analysis of a steady propagating shock is the solution for its jump conditions.

### 3. Shock Jump Conditions

Jump conditions for relativistic magnetized shocks were studied by de Hoffmann & Teller (1950). Pulsar wind nebulae and GRBs revived interest in relativistic shocks. The upstream medium is usually assumed to be cold in the sense that its enthalpy is much smaller than the rest-mass energy of the plasma. This condition is not satisfied, however, in the inner regions of GRB jets. Below we write down the general jump conditions for shocks propagating in a hot magnetized plasma filled with radiation, and show their solutions.

Consider a shock wave propagating in a sufficiently extended, optically thick medium. Far upstream and far downstream of the shock the plasma and radiation can be described as an ideal gas with isotropic pressure. The thermal energy density \( U \) and pressure are related by
\[
U = \frac{P}{\alpha - 1}, \tag{27}
\]
where \( \alpha = 4/3 \), as long as \( U \) is dominated by radiation. When formulating the jump conditions we will keep \( \alpha \) general, and specialize to \( \alpha = 4/3 \) in numerical solutions.

In the rest frame of the upstream (preshock) fluid, an observer will see the downstream (postshock) fluid approaching with velocity \( v_0 \). The shock front is perpendicular to \( v_0 \) and approaching with a higher velocity \( v_1 \) parallel to \( v_0 \). The
plasma carries a frozen magnetic field $B$ (measured in the fluid rest frame). We consider only magnetic fields perpendicular to the fluid velocity; a parallel magnetic field is anyway unchanged by the shock and hence does not affect the jump conditions.

### 3.1. Stress–Energy Tensor and Sound Speed

The stress–energy tensor of a hot magnetized flow with four-velocity $u^\mu = (\gamma c, \gamma \mathbf{v})$ is given by

$$ T^{\mu \nu} = \left( \tilde{\rho} c^2 + U + P \right) \frac{u^\mu u^\nu}{c^2} + g^{\mu \nu} P + \frac{1}{4 \pi} \left( F^{\alpha \beta} F_{\alpha \beta} - \frac{\delta^{\mu \nu}}{4} F^{\alpha \beta} F_{\alpha \beta} \right), \quad (28) $$

where $\tilde{\rho}$ is the proper rest-mass density of the baryons, $g^{\mu \nu}$ is the metric tensor of Minkowski spacetime, and $F^{\mu \nu}$ is the electromagnetic tensor. Its electric and magnetic components in the lab frame, $E$ and $B$, are related by $E + v \times B / c = 0$, as the plasma is a nearly ideal conductor. Using $\mathbf{v} \perp \mathbf{B}$ and $B = \gamma B$, the stress–energy tensor may be reduced to the ideal fluid form,

$$ T^{\mu \nu} = H_{\text{eff}} \frac{u^\mu u^\nu}{c^2} + g^{\mu \nu} P_{\text{eff}}, \quad (29) $$

with the effective relativistic enthalpy and pressure

$$ H_{\text{eff}} = \tilde{\rho} c^2 + U + P + \frac{B^2}{4 \pi}, \quad P_{\text{eff}} = P + \frac{B^2}{8 \pi}. \quad (30) $$

Before considering shocks, it is useful to examine sound waves in a uniform background that has $u_0^\mu = (c, 0, 0, 0)$, four-acceleration $u^\mu \nabla_\mu u^\nu = (0, \partial_\nu v, 0, 0)$, and compression $\nabla_\mu u^\nu = \partial_\nu v$. The linearized equations of motion $\nabla_\mu T^{\mu \nu} = 0$ give (for $\nu = t$ and $\nu = x$)

$$ \frac{\partial H_{\text{eff}}}{\partial t} + H_0 \frac{\partial v}{\partial x} - \frac{\partial P_{\text{eff}}}{\partial t} = 0, \quad (31) $$

$$ H_0 \frac{\partial v}{\partial t} + c_s^2 \frac{\partial P_{\text{eff}}}{\partial t} = 0. \quad (32) $$

These two equations can be reduced to the wave equation for $v$,

$$ \frac{\partial^2 v}{\partial t^2} - c_s^2 \frac{\partial^2 v}{\partial x^2} = 0, \quad (33) $$

where the wave speed $c_s$ is defined by

$$ c_s^2 = \frac{\partial^2 P_{\text{eff}}}{\partial t^2} - (H_{\text{eff}} - P_{\text{eff}}). \quad (34) $$

As the two main parameters of the fluid, it is convenient to use the dimensionless contributions of enthalpy and magnetic fields to the fluid inertia

$$ w \equiv \frac{U + P}{\tilde{\rho} c^2}, \quad \sigma \equiv \frac{B^2}{4 \pi \tilde{\rho} c^2}. \quad (35) $$

Then the wave speed defined in Equation (34) may be expressed as

$$ \frac{c_s^2}{c^2} = \frac{(\alpha - 1) w + \sigma}{1 + w + \sigma}. \quad (36) $$

This general expression reduces to familiar cases in four limits:

1. $\sigma \ll w \ll 1$: $c_s^2 = \alpha \rho \tilde{\rho}$ (nonrelativistic sound waves),
2. $w \ll \sigma \ll 1$: $c_s^2 = B^2 / 4 \pi \tilde{\rho}$ (nonrelativistic fast MHD modes in a cold plasma),
3. $w \gg \sigma, 1$: $c_s^2 = (\alpha - 1) c^2 = c_s^2 / 3$ (sound waves in a relativistic gas), and
4. $\sigma \gg w, 1$: $c_s = c$ (force-free limit of the MHD modes).

Internal supersonic motions $v_0 > c_s$ generate shocks, as discussed in detail in Section 2. In addition, shocks can form through nonlinear steepening of sound waves excited by a subsonic perturbation, $v_0 < c_s$ (Zeldovich & Raizer 1966). The steepening occurs because $c_s$ is slightly increased in the region compressed by the wave, so the crest of the wave (maximum $v > 0$ and maximum $\rho$) travels faster than the trough (minimum $v < 0$ and minimum $\rho$). Using Equation (36) one can verify that $dc_s / d\tilde{\rho} > 0$, i.e., compression indeed increases the local sound speed. The shock formed through sound-wave steepening propagates supersonically but has a subsonic velocity jump, i.e., it separates regions with a relative velocity $v_0 < c_s$. Such weak shocks are found among the solutions shown below, along with strong shocks formed by supersonic motions $v_0 > c_s$.

Formation of shocks through steepening of sound waves is inefficient in the relativistic regimes (3) and (4), because in this case $dc_s / d\tilde{\rho} \rightarrow 0$ ($c_s = c / \sqrt{3}$ or $c_s = c$ is constant in both cases). In the force-free limit ($w \ll \sigma \gg 1$), shock formation does not occur at all (Section 2.4). In a radiation-dominated medium ($\sigma \ll w \gg 1$), shocks can be launched by a supersonic motion $v_0 > c / \sqrt{3}$.

### 3.2. Jump Conditions

Jump conditions express the continuity of fluxes of energy, momentum, and baryon number in the rest frame of the shock front. The fluxes of energy and momentum are given by the stress–energy tensor $T^{\mu \nu}$ in Equation (29). The baryon flux is described by the four-vector

$$ F^\mu = \tilde{\rho} u^\mu. \quad (37) $$

The fluxes along the shock normal (the $x$-axis) are given by

$$ T^{x x} = \gamma P \tilde{\rho} c^2 (1 + w + \sigma), \quad (38) $$

$$ T^{x x} = P \tilde{\rho} c^2 (1 + w + \sigma) + P + \frac{B^2}{8 \pi}, \quad (39) $$

$$ F^x = P \tilde{\rho} c, \quad (40) $$

where $\rho = \gamma \beta$, and $P$ can be expressed in terms of $w$: $P = (1 - \alpha^{-1}) w \tilde{\rho} c^2$. Equating the fluxes upstream (index “$u$”) and downstream (index “$d$”) one obtains the relations

$$ \frac{T^{x x}}{c F^x} = \gamma_d (1 + w_d + \sigma_d) = \gamma_u (1 + w_u + \sigma_u), \quad (41) $$

$$ \frac{T^{x x}}{c F^x} = P_d (1 + w_d + \sigma_d) + \left( 1 - \frac{1}{\alpha} \right) \frac{w_d}{P_d} + \frac{\sigma_d}{2 P_d} \Rightarrow P_u (1 + w_u + \sigma_u) + \left( 1 - \frac{1}{\alpha} \right) \frac{w_u}{P_u} + \frac{\sigma_u}{2 P_u}. \quad (42) $$

Given the upstream parameters $p_u$, $w_u$, $\sigma_u$, and taking into account that $p_u \sigma_u = p_d \sigma_d$ (implied by the flux freezing condition $B \propto \tilde{\rho}$), one can solve Equations (41) and (42) for the two unknowns $p_d$ and $w_d$. 

7
Typically, the upstream velocity relative to the downstream, $v_0$, is a given in the shock problem. Therefore, we chose $g = p v_0 c_0$ as an independent parameter instead of $p u$. The upstream momentum in the shock frame, $p_u$, is related to the upstream momentum measured in the downstream frame, $p_0$, by the Lorentz transformation between the two frames,

$$g = p u = g_v (p_0 + \beta_v v_0).$$  \hfill (43)

For a given $p_0$, a trial $p_d$ determines $p_u (p_d)$, and the solution of Equations (41) and (42) (which is obtained numerically) yields $p_d$ and $p_u$ together with $w_d$.

The solutions are shown for $p_0 = 1$ and 10 in Figures 4 and 5, assuming $\alpha = 4/3$ (pressure is dominated by radiation). Figure 4 shows the compression ratio $\xi = p_d / p_0$ for shocks with $p_0 = 1$ (upper panel) and $p_0 = 10$ (lower panel). This ratio depends on the upstream enthalpy $w_u$ and magnetization $s_u$; the dependence is shown using contours on the $s_u-w_u$ plane. The blue region is the same as in Figure 4. Red dots highlight the region where $w_d < 2 s_u$. In this region, the shock is expected to have a strong collisionless jump where the velocity changes on a scale comparable to the gyroradius of the plasma particles.

Figure 4. Compression ratio $\xi = p_d / p_0$ in a relativistic shock with $p_0 = 1$ (upper panel) and $p_0 = 10$ (lower panel). The compression $\xi$ depends on the upstream enthalpy $w_u$ and magnetization $s_u$. Solid curves show contours of the function $\xi(s_u, w_u)$. The region where the shock is weak, $v_0 < c_s$, is dotted in blue and bounded by the dashed curve.

Figure 5. Ratio of the enthalpy and magnetization in the downstream, $w_d / s_d$, for shocks with $p_0 = 1$ (upper panel) and $p_0 = 10$ (lower panel). This ratio depends on the upstream enthalpy $w_u$ and magnetization $s_u$; the dependence is shown using contours on the $s_u-w_u$ plane. The blue region is the same as in Figure 4. Red dots highlight the region where $w_d < 2 s_u$. In this region, the shock is expected to have a strong collisionless jump where the velocity changes on a scale comparable to the gyroradius of the plasma particles.
3.3. Collisionless Shocks

The jump conditions do not describe the structure or dissipation mechanism of the shock front. However, they allow one to evaluate the region in the parameter space where dissipation must be mainly collisionless.

We expect the shock to be mainly mediated by collective electromagnetic fields when the downstream enthalpy \( w_d \) is dominated by the compressed magnetic field \( \sigma_d \). Then radiation cannot control the shock structure, because its pressure is below the ram pressure of the shock. In particular, in the limit of \( w_d \ll \sigma_d \) the downstream plasma can be approximated as a cold magnetized medium with negligible heat, so radiation has no effect on the upstream deceleration and the shock velocity profile; the profile inevitably steepens so that the entire velocity jump occurs in a thin layer on the collisionless plasma scale (gyroradius).

In contrast, in shocks with significant ratio \( w_d / \sigma_d \), the diffusion of the postshock radiation into the upstream region creates a precursor that changes the upstream velocity and reduces the amplitude of the collisionless jump. The resulting structure may be described as a collisionless shock with a radiation precursor or, equivalently, an RMS with a collisionless subshock. In the limit of \( w_d \gg \sigma_d \) the subshock becomes weak or nonexistent. Section 4 below demonstrates this fact for shocks in relativistic gas, \( w \gg 1 \), and Section 5 will show the shock structure for a moderate \( w = 0.1 \) with and without a significant magnetic field.

The transition between the two dissipation regimes—mainly mediated by collective electromagnetic fields and mainly mediated by radiation—occurs at \( w_d / \sigma_d \approx 2 \) (the shock structure in this transition region will be calculated in Section 5). The region where a strong collisionless jump is expected \( (w_d / \sigma_d \lesssim 2) \) is highlighted in red in Figure 5. We also show (in blue) the region where the shock is weak \( (w_0 < c_s) \) and could form only through steepening of sound waves. As discussed in Section 3.1, such shocks do not easily form in a relativistic fluid \( (w > 1 \text{ or } \sigma > 1) \) since steepening takes a long time, typically longer than the expansion time of the outflow. The region between the two curves \( w_d = 2 \sigma_d \) and \( w_0 = c_s \) is where strong collisionless shocks occur. When \( w_d \ll \sigma_d \) only a small fraction of the upstream kinetic energy is dissipated in the shock, and most of it ends up stored in the compressed magnetic field. Therefore, the strongest collisionless dissipation is expected if \( w_d \sim \sigma_d \).

For a medium with a given enthalpy \( w_0 \), one can define a characteristic magnetization \( \sigma_0 = \sigma_e \) such that a shock propagating in the medium will have \( w_d / \sigma_d = 2 \) (the boundary of the red dotted region in Figure 5). The magnetization \( \sigma_0 \) depends on the shock strength \( p_0 \) (upstream momentum measured in the downstream frame). This dependence is shown in Figure 6. In the ultrarelativistic limit \( p_0 \gg 1 \), we find that \( \sigma_0 \) does not depend on \( p_0 \) (both \( w_0 \) and \( \sigma_d \) scale as \( p_0^2 \), so their ratio does not depend on \( p_0 \)). For nonrelativistic shocks \( p_0 \ll 1 \) with a cold upstream region \( (w_0 = 0) \), \( \sigma_0 \) scales as \( p_0^3 \). This is because weakly magnetized shocks have downstream enthalpy \( w_d \propto p_0^2 \) while \( \sigma_d \propto p_0 \).

At large \( \sigma > 1 \) and \( \sigma > w \), the effective sound speed \( c_s \) approaches the speed of light (see Equation (36)). Shocks easily form if the internal compressive motions are supersonic, i.e., their Lorentz factors \( \gamma \) exceed \( \sigma^{1/2} \). The dissipation in magnetically dominated shocks occurs in a microscopically thin, collisionless shock front. Dissipation is reduced in this regime, because a large fraction of shock energy goes into the compressed magnetic field, but dissipation can still be significant. For example, a shock with Lorentz factor \( \gamma = 4 \) propagating in a medium with upstream enthalpy \( w = 1 \) and \( \sigma = 10 \) has the downstream enthalpy \( w_d \approx 6.8 \) and \( U_\text{rad} \approx 0.12 U_{\text{th}} \). In this example, the shock compression ratio is \( 3 < 8.3 \), and adiabatic compression would imply the amplification of \( w \) by only a factor of \( \xi^{1/3} \approx 2 \), well below 6.8. We note also that the downstream enthalpy \( w_d \) is mainly determined by the upstream enthalpy \( w_0 \) and the amplitude of the shock; it depends weakly on \( \sigma_0 \).

4. Shocks in Photon Gas

In GRB explosion models, sub-photospheric shocks begin to form at early stages, when the rest mass of the jet is still dominated by radiation, before the jet accelerates to its asymptotic Lorentz factor. This section examines the structure of shock waves in this regime, neglecting magnetic fields.

In essence, we deal here with shocks in the gas of photons, because the plasma inertia is negligible. The role of plasma is to provide opacity and thus to couple the photons into a single fluid, with a small mean free path. The plasma particles may be viewed as passive “markers” following the motion of the photon gas.

**4.1. Jump Conditions**

Far upstream and far downstream of the shock, the radiation can be described as an ideal fluid with isotropic pressure \( P \) and the stress–energy tensor

\[
T^{\mu\nu} = 4P \delta^{\mu\nu} + Pg^{\mu\nu}. \tag{44}
\]

where we have used the equation of state \( U = 3P \). Jump conditions express the continuity of \( T^{tt} \) (flux of energy) and \( T^{xx} \).
(flux of momentum) in the rest frame of the shock,
\[ 4P_d \gamma_d^2 \beta_d = 4P_u \gamma_u^2 \beta_u, \]  
\[ 4P_d \gamma_d^2 \beta_d^2 + P_d = 4P_u \gamma_u^2 \beta_u^2 + P_u, \]  
where subscript “u” stands for upstream and “d” for downstream; pressure is measured in the fluid frame, and velocity is measured in the shock frame. Dividing Equations (45) and (46), one finds that \( \beta_u \) and \( \beta_d \) satisfy the condition
\[ g(\beta_u) = g(\beta_d), \quad g(\beta) = \frac{1 + 3\beta^2}{\beta}. \]  
Rewriting the definition of \( g \) as
\[ 3\beta^2 - g(\beta) + 1 = 0, \]
on one can view \( \beta_u \) and \( \beta_d \) as the two roots of the quadratic equation, and hence they are related by
\[ \beta_u \beta_d = \frac{1}{3}. \]  
Since \( \beta_u > \beta_d \), one concludes that \( \beta_u > 3^{-1/2} \). This condition merely states that the shock moves supersonically relative to the upstream plasma (recall that the sound speed is \( c_0 = 3^{-1/2}c \)). Using the relation (49) and Equation (45) or (46), one finds the pressure jump across the shock,
\[ \frac{P_d}{P_u} = 3\gamma_d^2 \left( \beta_u - \frac{1}{3} \right). \]  
The shock compresses the volume measured in the fluid frame by the factor
\[ \xi = \frac{\gamma_u \beta_u}{\gamma_d \beta_d} = \gamma_u \beta_u (9\beta_u^2 - 1)^{1/2}, \]  
which also gives the relation
\[ \frac{P_d}{P_u} = \frac{\xi^2}{3\beta_u^2}. \]  
For ultrarelativistic shocks, \( \beta_u \to 1 \), the jump conditions simplify to \( \beta_d = 1/3 \) and \( P_d/P_u = (8/3)^{\gamma_u^2} = \xi^2/3 \).

4.2. Evolution Equation for the Photon Gas

The formation of shocks in the gas of photons can be simulated numerically. It is convenient to think of this problem as a radiative transfer problem for the bolometric intensity of radiation \( I \). Since the stress–energy tensor is dominated by radiation, the plasma is effectively massless and its velocity \( \beta \) is controlled by the “force-free” condition: \( \beta \) equals the equilibrium value such that the radiation flux in the fluid frame vanishes (zero flux implies zero force applied by radiation to the plasma). This condition leads to a well-defined radiative transfer problem (Beloborodov 1999). It has a simple solution for steady spherically symmetric relativistic outflows (Beloborodov 2011). Here we are interested in shock formation in variable outflows, so the problem is time-dependent.

The shock is thin and locally flat (in the \( y-z \) plane), and we can study its formation in the plane-parallel geometry. Then the bolometric intensity is described by the function \( I(t, x, \mu) \) where \( \mu = \cos \theta \) and \( \theta \) is the photon angle with respect to the \( x \)-axis.

The stress–energy tensor of radiation is determined by the moments of the intensity,
\[ I_k(t, x, \mu) = \frac{1}{2} \int_{-1}^{1} I(t, x, \mu) \mu^k \, d\mu. \]
In particular, \( T^{uu} = 4\pi \rho_0, \ T^{ux} = 4\pi \beta_0 \), and \( T^{xx} = 4\pi I_0 \). The free-force condition reads \( \ddot{I}_0 = 0 \) in the fluid frame, and the transformation of \( T^{uu} \) from the lab frame to the fluid frame gives the quadratic equation for velocity,
\[ \ddot{I}_0 = \gamma^2 [-\beta(\dot{I}_0 + \dot{L}) + (1 + \beta^2)I_0] = 0, \]
\[ \Rightarrow \beta = \gamma (\gamma^2 - (\gamma^2 - 2)^{1/2}), \quad \gamma \equiv \frac{\dot{I}_0 + \dot{L}}{I_0}. \]

The evolution of intensity is described by the transfer equation,
\[ \frac{1}{c} \frac{\partial I}{\partial t} = -\mu \frac{\partial I}{\partial x} + n \sigma_\tau (1 - \beta \mu)(S - I), \]
where \( n \) is the number density of electrons/positrons measured in the lab frame, and \( S \) is the source function. In the simplest case of isotropic scattering, \( S \) is given by (Beloborodov 1999)
\[ S(\mu) = \frac{\dot{I}_0 - \dot{\beta}_0}{\gamma^2 (1 - \beta \mu)^2}. \]

4.3. Numerical Solution

Equation (56) supplemented with the equilibrium velocity condition (Equation (55)) can be solved numerically. A sample solution is shown in Figure 7. In this example, the initial state is given by Equation (2) with \( \rho_{p0} = 2 \) and \( L = 50 \). The initial plasma density is uniform in the lab frame, \( \rho(0, x) = \rho_0 = \text{const.} \), and the unit length in \( x \) corresponds to a slab of unit Thomson optical depth. The initial radiation density measured in the fluid frame is uniform in the lab frame, \( U(0, x) = U_0 = \text{const.} \).

One can see the compression of the converging supersonic flow and the formation of a pair of shocks symmetric about \( x = 0 \), as described in Section 2. As the two shocks continue to propagate, the downstream fluid comes nearly to rest in the lab frame. The upstream velocity relative to the downstream fluid, \( \beta_0 \), is related to the upstream and downstream velocities measured in the shock frame by
\[ \beta_0 = \frac{\beta_u - \beta_d}{1 - \beta_u \beta_d}. \]

Using \( \beta_d \beta_u = 1/3 \) (Section 4.1), one finds
\[ \beta_u = \frac{1}{3} (\beta_0 + (\beta_0^2 + 3)^{1/2}). \]

Once the shock wave is established, its structure becomes independent of the details of the initial conditions. The shock has only one parameter: \( \beta_0 \) or \( \rho_{p0} = \gamma_0 \beta_0 \). Figure 8 shows the structure of a shock wave obtained with \( \rho_{p0} = 2 \). It is shown as a function of the optical depth measured in the \( x \) direction. Then the result is independent of the plasma density, so the obtained solution is unique. Using \( \beta_0 = 2^{-1/2} \) and \( \beta_u = 0.7903 \), which correspond to \( \rho_0 = 2 \), one finds from Equations (51) and (52) the ratio of downstream and upstream pressures \( P_d/P_u = U_d/U_u = 23.3 \). This asymptotic value of \( U \) is observed in Figure 8.
For comparison, Figure 8 also shows the momentum profile for a shock with \( \gamma = \frac{40}{\gamma} \), obtained from a similar time-dependent simulation. When rescaled by the factor of 2, the momentum profile is the same as for \( \gamma = \frac{20}{\gamma} \).

5. Monte-Carlo Simulations of Shocks

Time-dependent simulations may also be employed to study shocks in plasma with significant rest mass and magnetic fields. In contrast to the photon gas studied in Section 4, now the radiative transfer equation cannot be closed by the force-free condition \( \tilde{I} = 0 \). Instead, the fluid acceleration must be calculated together with the radiative transfer. Another complication is the need to follow the evolution of the radiation spectrum, which develops a hard tail extending above \( \sim m_{\text{e}}c^2 \) inside the shock front; then the scattering cross section is changed by the electron recoil.

Below we solve this problem using a direct numerical experiment that follows individual photons and their interactions with the plasma, so that the transfer of momentum and energy is described on a microscopic level. In the initial state, the flow has a smooth velocity profile described by Equation (2) and carries thermal radiation, which is isotropic in the fluid frame. The flow is opaque and supersonic, which leads to the formation of a pair of shocks propagating in the \( \pm x \) directions, as described in Section 2.

The flow has two interacting components.

1. Magnetized plasma. The plasma is assumed to carry a transverse magnetic field, which provides strong coupling between all charged particles, so their dynamics along the \( x \)-axis is well described as a single-fluid motion.\(^3\) In the

---

\(^3\) For any realistic magnetic field the Larmor radii of ions and electrons are microscopic, many orders of magnitude smaller than the photon free path. Therefore, we assume strong magnetic coupling of the plasma particles even in the weakly magnetized model that is formally labeled as \( \sigma = 0 \).
numerical simulations we use a Lagrangian grid moving together with the fluid of charged particles: the fluid is discretized into $N \sim 10^4$ shells of equal rest mass $m$ and a small scattering optical depth. Besides mass, each shell is characterized by the magnetic flux frozen in it, internal thermal energy, and total pressure. The magnetic flux remains constant while thermodynamic quantities may change as the shell contracts (or expands) and interacts with radiation. Thermal conductivity of the plasma in the $x$ direction is suppressed by the transverse magnetic field and neglected.

(2) Radiation. Radiation is represented by $\sim 10^8$ photons, which are followed individually. The photons migrate through the plasma shells and occasionally scatter off a thermal electron. The scattering is followed using the Monte-Carlo technique, with the exact Klein–Nishina differential cross section and assuming that the thermal electrons are isotropic in the fluid frame. The electrons are assumed to have a Maxwellian distribution with a self-consistently calculated temperature.

A detailed description of the numerical method will be given in a forthcoming paper. A possible alternative to the Monte-Carlo method is the solution of the transfer equation for the radiation intensity, as in Section 4 but now including fluid inertia. A similar approach was taken in the recent work by Ohsuga & Takahashi (2016), who simulated the evolution of bolometric intensity assuming Thomson scattering. The Thomson approximation may be sufficient only for the RMS with negligible fluid inertia $\tilde{p}c^2 \leq U_{\text{rad}}$. If $\tilde{p}c^2 > U_{\text{rad}}$, strong bulk Comptonization develops in the RMS and, if treated in the Thomson approximation, leads to runaway in photon energy (Blandford & Payne 1981). Scattering with substantial electron recoil, in the Klein–Nishina regime, becomes inevitable and limits the growth of photon energy. The electron recoil is also essential in maintaining heat exchange between radiation and plasma, even when the photon energies (and the plasma temperature) are well below $m_e c^2$.

### 5.1. Sample Models

In our two sample simulations the initial flow has dimensionless enthalpy $w = 0.1$ and magnetization $\sigma = 0$ (Model A) or $\sigma = 0.1$ (Model B). The initial average photon energy in the fluid frame is everywhere $3kT \approx 10^{-2}m_e c^2$. In both simulations we observed how the compressive wave with amplitude $p_{\text{max}} = 2$ steepened and formed a pair of shocks at time $t \approx L/c$. We chose a sufficiently large optical depth of the steepening region $\tau_L = \tau_0 n L = 20$ and observed how the magnetic field and the trapped radiation were advected toward $x = 0$, building up a strong pressure maximum. This launched the RMS, as described in Section 2, and the two symmetric shocks continued to propagate away from $x = 0$. The amplitude of the shocks $p_0$ grows slowly as they propagate toward the asymptotic momentum of the converging flow $p_{\text{max}} = 2$.

Figure 9 shows one of the two symmetric shocks at $t \sim 3L/c$ in Model A (upper panel) and Model B (lower panel). By this time the shock has crossed a Thomson optical depth $\tau_T \approx 60$ from its formation site, and the upstream momentum has reached $p_0 \approx 1.6$. The shock structure is steady and propagating relative to the downstream fluid with speed $v \approx 0.3v_0$. The shock exhibits the jump conditions calculated in Section 3. In particular, the shock compression ratio in Model B is $\xi = \tilde{p}_T/\tilde{p}_u \approx 6.8$. The jump conditions also give a moderate ratio $\sigma_T/\nu_T \approx 0.6$; it turns out to be sufficient to form a strong collisionless subshock. The simulation confirms the expectation from Section 3 that the flow magnetization leads to a strong collisionless subshock in the RMS. The subshock becomes weak if the magnetization is reduced and disappears in Model A with $\sigma = 0$.

The observed shock structure in Model B ($\sigma = 0.1$) may be summarized as follows. The momentum profile $p(\tau_T)$ is shallow toward the upstream; this part is shaped by radiation pressure that gradually decelerates the upstream on a scale comparable to the photon mean free path $L_\text{ph}$. The profile drops steeply toward the downstream and has a kink connecting to
Bromberg 2008): a fraction of photons cross the shock back and forth multiple times, with the energy boost \( \sim \gamma_0^2 \) in every cycle, similar to Fermi diffusive acceleration. As a result, the photon spectrum extends somewhat above \( m_e c^2 \) in the fluid frame. Further energy growth is hindered by downscattering due to the strong electron recoil (and also by photon conversion to \( e^\pm \) pairs, see below). At large optical depths downstream of the shock, the multiple downscattering of high-energy photons drives the spectrum toward a Wien shape in Compton equilibrium with the electrons.

5.2. Pair Creation

Let us consider the weakly magnetized RMS, with no collisionless subshock. The ability of bulk Comptonization to generate photons with energies \( E > m_e c^2 = 511 \text{ keV} \) in the fluid frame implies a significant rate of \( e^\pm \) pair creation due to the reaction \( \gamma + \gamma \rightarrow e^+ + e^- \). This reaction was not included in our Monte-Carlo simulations, and below we discuss its effect on the RMS structure.

The rate of pair creation can be estimated by noting that the absorbed MeV photons are replenished at a rate \( n_1 / t_1 \) where

\[
\frac{n_1}{t_1} = \frac{dn_1}{d\ln E} E \approx \frac{1}{m_e c^2} \frac{d\ln \epsilon}{d\ln E},
\]

and \( t_1 \sim (3 - 10) \ell_{ph} / c \) is the time it takes a 0.5 MeV photon to double its energy through bulk Comptonization. The radiation spectrum inside the RMS shows \( n_1 \sim 10^{-2} n_\gamma \), where \( n_\gamma \) is the total photon density. Therefore, one can roughly estimate the pair production rate as

\[
\dot{n}_\pm \sim 10^{-3} \frac{c n_\gamma}{\ell_{ph}^3}.
\]

Combining this with the characteristic timescale of the shock propagation, \( \ell_{ph} / c \), one obtains an order-of-magnitude estimate for the pair density in the RMS, \( n_\pm \sim 10^{-2} n_\gamma \). The corresponding number of \( e^\pm \) created per proton is

\[
Z_\pm \sim 10^{-3} \frac{n_\gamma}{n} = 10^2 \left( \frac{n_\gamma / n}{10^5} \right).
\]

where \( n_\gamma / n \sim 10^5 \) is a typical photon-to-baryon ratio in GRBs.

This estimate neglects two effects, which somewhat reduce \( Z_\pm \).

1. Equation (61) assumes that the photons reaching \( \sim 1 \text{ MeV} \) are quickly absorbed in \( \gamma-\gamma \) collisions, neglecting their finite free path. The MeV photons collide with each other, because they are near the threshold for \( \gamma-\gamma \) reaction, \( 2m_e c^2 \approx 1 \text{ MeV} \). Their free path \( \ell_{\gamma\gamma} \) may be estimated using the cross section for \( \gamma-\gamma \) collision \( \sigma_{\gamma\gamma} \approx 0.1 \sigma_T \). Then one finds \( \ell_{\gamma\gamma} \sim 10(\sigma_T n_\gamma)^{-1} \sim 10^3 Z_{\pm} n_\gamma / n_\gamma \ell_{ph} \), and

\[
\frac{\ell_{\gamma\gamma}}{\ell_{ph}} \sim \frac{Z_{\pm}}{10^2} \left( \frac{n_\gamma / n}{10^5} \right)^{-1} \sim 1.
\]

Hence a large fraction of \( e^\pm \) creation occurs inside the RMS, as assumed in Equation (61), and so the finite \( \ell_{\gamma\gamma} \) is not a big change.

2. Equation (61) neglects the effect of \( e^\pm \) annihilation. The annihilation rate \( \dot{n}_{\text{ann}} = (3/8) \sigma_T c n_\gamma n_\gamma \) implies a positron lifetime \( \tau_{\text{ann}} = (8/3)(\sigma_T n_\gamma)^{-1} \). The positrons are advected by the flow with a mildly relativistic speed and annihilate over a distance that corresponds to Thomson optical depth \( \sim 1 \), comparable to the RMS thickness. Hence annihilation has a moderate reducing effect on \( Z_\pm \) inside the RMS.

The Astrophysical Journal, 838:125 (21pp), 2017 April 1

Beloborodov
These estimates show that $Z_{\pm}$ inside the RMS can approach $\sim 10^2$. The exact value can be obtained from detailed numerical simulations and will depend on the upstream temperature, strength of the shock $\gamma_0$, and $n_{p}/n$. $Z_{\pm}$ peaks inside the RMS and decreases with distance downstream, because of $e^\pm$ annihilation.

The mildly relativistic pairs created are almost immediately cooled by inverse Compton (IC) emission and Coulomb collisions with the background plasma. The corresponding energy loss rates of an electron with Lorentz factor $\gamma = (1 - \beta^2)^{-1/2}$ are given by (e.g., Ginzburg & Syrovatskii 1964)

$$E_{\text{IC}} \approx \frac{4}{3} \sigma_T U_{\text{rad}} \gamma^2 \beta^2,$$

$$E_{\text{Coul}} \approx \frac{3}{2} \ln \Lambda \sigma_T n_{\pm} m_e c^3,$$

where $\ln \Lambda = \ln(m_e c^2/\hbar \omega_{\text{pl}}) \approx 20$ is a Coulomb logarithm ($\omega_{\text{pl}}$ is the Langmuir plasma frequency). Their ratio may be written as

$$\frac{E_{\text{IC}}}{E_{\text{Coul}}} \approx \frac{10^2}{Z_{\pm}\gamma^2 \beta^2},$$

where $Z_{\pm} = n_{\pm}/n$ is the number of $e^\pm$ per proton. For the typical parameters, the Coulomb losses dominate when $\gamma \beta \lesssim 1$. Thus the injected mildly relativistic particle shares a significant fraction of its energy with the background plasma and deposits momentum $\sim m_e c$. The estimated rate of $\gamma-\gamma$ absorption then implies the force $f_{\text{abs}} \sim n_{\pm} m_e c$ exerted on the plasma. This additional force is modest compared with the radiation pressure force $f_{\text{sc}} \sim n_{\pm} c^2/\rho c$,

where $n_{\pm} = n \gamma = n_{p} c/\ell_{\text{ph}}$ is the scattering rate and $\rho_{p} m_e c^2 \sim 10$ keV is the average photon energy. Using the above estimates one finds $f_{\text{abs}}/f_{\text{sc}} \sim n_{\pm}/n \rho c \lesssim 0.1$.

Pair creation increases the local scattering opacity $\kappa(x) = Z_{\pm}(x) \sigma_{\text{T}} m_p$ by the factor $Z_{\pm}$. In essence, an additional $e^\pm$ “screen” is created between the upstream and downstream. The thickness of a relativistic RMS is always a few times $\ell_{\text{ph}}$, which is a few times $\tanh^{-1} Z_{\pm}$. For simplicity, the Monte-Carlo simulations presented in Section 5.1 assumed $\kappa = \text{const}$. This assumption is not important as long as (1) the shock structure is steady and (2) the velocity and density profiles of the shock wave are viewed as functions of the scattering optical depth (as in Figure 9) rather than the spatial coordinate $x$.

In the plane-parallel approximation (valid when the RMS thickness is much smaller than its radius) the actual length $\ell_{\text{ph}}$ does not matter, and the optical depth is the natural coordinate.

### 6. Collisionless (Sub)shock Heating

In this section, $v_0 = \beta_{0} c$ refers to the strength of the collisionless shock, which may be embedded in a stronger RMS. The shock thickness is microscopic, comparable to the ion Larmor radius.

The heat generated by the shock is initially given to the plasma particles and then converted to radiation, at some distance downstream of the shock. In particular, the postshock ions receive thermal speeds $v_{\text{th}} \sim v_0$, as the shock thermalizes their upstream bulk speed $v_0$. Numerical simulations of collisionless electron–ion shocks show that fast collective processes help the ions to promptly pass a fraction $f_e \sim 0.3–0.5$ of their energy to the electrons, and both form Maxwellian distributions (Sironi & Spitkovsky 2011).

The fraction $f_e$ is less studied for the most interesting shocks in pair-loaded plasma, with $1 < Z_{\pm} < m_i/m_e$, where $m_i$ is the ion mass. The existing work on pair-loaded shocks with transverse magnetic fields focused on the ultrarelativistic regime, with application to pulsar wind nebulae. It was found that such shocks are capable of positron acceleration through the synchrotron maser instability of gyrating ions (Hoshino et al. 1992; Amato & Arons 2006; Stockem et al. 2012). It is unclear whether particle acceleration may be efficient in moderately relativistic internal shocks in GRBs.

Below we discuss how the postshock plasma radiates its energy, assuming a two-temperature state. The electron and ion temperatures immediately behind the shock are controlled by the parameter $f_e$. The thermal Lorentz factor of the ions, $\gamma_{\text{th}}$, is related to the upstream Lorentz factor $\gamma_0 = (1 - \beta^2)^{-1/2}$ by

$$\gamma_{\text{th}} - 1 = (1 - f_e)(\gamma_0 - 1) \quad \text{(ions)}. \tag{66}$$

The electrons are heated to much higher Lorentz factors,

$$\gamma_{\text{th},e} = f_e (\gamma_0 - 1) \frac{m_i}{Z_{\pm} m_e} \gg 1 \quad \text{(electrons)}, \tag{67}$$

where $Z_{\pm} = n_{\pm}/n$ is the number of $e^\pm$ pairs per proton upstream of the collisionless shock. We will first discuss electron cooling and then ion cooling.

#### 6.1. Electron Cooling

The suddenly heated electrons lose their energy to IC scattering on a short timescale,

$$t_{\text{IC}} = \frac{3 m_e c}{4 \sigma_T f_{\text{KN}} U_{\text{rad}} \gamma_{\text{th},e}}, \tag{68}$$

where the factor $f_{\text{KN}} < 1$ describes the Klein–Nishina correction to the Compton cooling rate. Even accounting for pair creation, the photon number in GRB jets exceeds the electron number by a large factor (3–5 orders of magnitude when the jet Lorentz factor saturates). Therefore, the electron cooling time $t_{\text{IC}}$ is much shorter than the free-path time of photons to scattering, $t_{\text{IC}} \ll \ell_{\text{ph}}/c$.

The electrons are also cooled by synchrotron losses on the timescale

$$t_{\text{syn}} = \frac{3 m_e c}{4 \sigma_T U_{\text{th},e} \gamma_{\text{th},e}}. \tag{69}$$

It is shorter than $t_{\text{IC}}$ if $U_{\text{th}} > f_{\text{KN}} U_{\text{rad}}$.

The maximum energy of IC photons is determined by $\gamma_{\text{th},e}$, which depends on $Z_{\pm}$ (Equation (67)). IC photons with energies $E_{\text{IC}} > 1$ MeV are processed through $e^\pm$ cascade into secondary $e^\pm$ pairs and eventually into photons of energy $\lesssim 1$ MeV, which are capable of escaping $\gamma-\gamma$ absorption (Svensson 1987). As a result, most $e^\pm$ pairs are created in collisions between photons of energy $E_{\text{IC}} \sim 1$ MeV, not much above the $2m_e c^2$ threshold. The situation resembles that described in Section 5.2.

The multiplicity of pairs created per shock-heated electron, $M_{\pm}$, is maximum if synchrotron losses are small, $U_{\text{th}} \ll f_{\text{KN}} U_{\text{rad}}$.

---

4 Photos of energy $E_{\text{IC}}$ are absorbed in collisions with photons of energies $E_{\gamma} \gtrsim 2m_e c^2/E_{\text{IC}}$ with cross section $\sigma_{\gamma} \sim 0.1 \sigma_T$. The mean free path to absorption is $\ell_{\gamma} = (\sigma_T n_\gamma)^{-1}$ where $n_\gamma$ is the target photon density. For instance, 100 MeV photons are absorbed by the Wien peak of the radiation spectrum ($E_{\gamma} \sim 10$ keV in the fluid frame) and then the $\ell_{\gamma} \sim 10 (\sigma_T n_\gamma)^{-1}$ is tiny.
then it approaches $\mathcal{M}_{\pm} \sim 0.2\gamma_{\text{th,e}}$. If the magnetic energy is comparable to the radiation density, synchrotron losses become dominant. This significantly reduces the pair yield $Y = \mathcal{M}_{\pm}/\gamma_{\text{th,e}}$ (see Table 1 in Vurm et al. 2011).

The population of MeV photons with peak density $n_{\text{MeV}}$ spreads from the collisionless shock over the characteristic distance of their self-destruction in $\gamma\gamma$ collisions,

$$\ell_{\gamma\gamma} \sim (\sigma_{\gamma\gamma} n_{\text{MeV}})^{-1}.$$  \hfill (70)

The number of MeV photons emitted per shock-heated electron approximately equals the number of pairs they produce, i.e., $\mathcal{M}_{\text{MeV}} \sim \mathcal{M}_{\pm}$.

The $e^\pm$ density in the pair-creation region of size $\sim \ell_{\gamma\gamma}$ is somewhat reduced by the annihilation reaction with rate $n_{\text{ann}} = (3/8)\sigma_{\gamma\gamma} n n$. The downstream evolution of the positron density $n_+ \sim n_{n\gamma\gamma}$ is described by the equation

$$v_d \frac{d n_+}{d x} = \frac{3}{8} \sigma_{\gamma\gamma} n_+ n,$$  \hfill (71)

Here $\sigma_{\gamma\gamma} \sim \sigma_{\gamma\gamma} c n_{\text{MeV}}^2$ is the pair creation rate, and $v_d \approx v_0/3$ is the velocity of the downstream fluid relative to the collisionless shock.\(^5\) The advection rate across the pair-creation zone $v_d \frac{d n_+}{d x} \sim n_+ v_d/\ell_{\gamma\gamma}$ is comparable to or smaller than the terms on the right side of Equation (71). Therefore, the pair density is not far from the annihilation balance, $(3/8)\sigma_{\gamma\gamma} c n_+ n \sim \sigma_{\gamma\gamma} c n_{\text{MeV}}$. This is still consistent with the crude estimate for the positron density $n_+ \sim n_{n\text{MeV}}$, which determines the Thomson optical depth of the pair creation zone,

$$n_+ \sim n_{\text{MeV}}, \quad \tau_\gamma \sim \sigma_{\gamma\gamma} \ell_{\gamma\gamma} \sim 2.$$  \hfill (72)

The production of MeV photons that convert to pairs is different from the RMS bulk Comptonization discussed in Section 5. However, it gives an optical depth comparable to that of an RMS.

### 6.2. Pair Breeding Upstream of Collisionless Shocks

The strong electron heating in collisionless shocks results in the MeV emission downstream of the shock; however, some of the MeV photons can convert to pairs in the upstream. This leads to a special type of self-sustained breeding of pairs.

The number of $e^+$ and $e^-$ created per shock-heated electron may be written as $\mathcal{M}_\pm = \mathcal{M}_e + \mathcal{M}_d$—the sum of $e^\pm$ created upstream and downstream. Most of the pair-creating MeV photons are emitted relatively close to the collisionless shock, at a distance smaller than their absorption free-path $\ell_{\gamma\gamma}$.\(^6\) Therefore, IC photons emitted toward the upstream can overtake the shock and convert to pairs ahead of it.

5. The relation $v_d \approx v_0/3$ holds for plasma-mediated shocks with any amplitude $v_0$, relativistic or nonrelativistic, as long as the shock is strong ($v_0$ is much greater than the upstream thermal speed). The relation $v_d = v_0/3$ is derived analytically from the jump conditions in the “monoenergetic gas” approximation (Beloborodov & Uhm 2006).

6. In a GRB jet, the average photon energy in the fluid frame is comparable to 10 keV. Most of the pair-creating IC photons are generated in the Thomson regime, $E_{\text{IC}} \sim \gamma_e^2 \times 10 \text{keV}$, by $e^\pm$ with $\gamma_e \lesssim 30$ over length $L_{\text{IC}} \approx \nu_{\text{th,e}} \approx 0.1/m_e c^2/\nu_{\text{th,e}}$, see Equation (68). The length $L_{\text{IC}}$ is much shorter than $\ell_{\gamma\gamma} (\text{E}_{\text{IC}})$. If pair creation involves a cascade with multiple generations, many of the electrons with $\gamma_e < 30$ would be secondary and injected over an extended length due to the intermediate steps of $\gamma\gamma$ absorption. However, even in this case the additional steps are relatively short, because $\ell_{\gamma\gamma} (100 \text{MeV}) \ll \ell_{\gamma\gamma} (10 \text{MeV}) \ll \ell_{\gamma\gamma} (1 \text{MeV})$.

A rough estimate for $\mathcal{M}_u/\mathcal{M}_\pm$ may be obtained by assuming that the IC photons are emitted isotropically in the downstream fluid frame and that photons emitted with angles $\cos \theta > \beta_d$ (catching up with the shock) have a sufficiently long $\ell_{\gamma\gamma}$ to overtake the shock. Then,

$$\mathcal{M}_u \sim \frac{1 - \beta_d^2}{2} \mathcal{M}_{\pm}.$$  \hfill (73)

A more accurate value for $\mathcal{M}_u$ may be found from Monte-Carlo transfer simulations. It will be somewhat below the estimate given in Equation (73).

$\mathcal{M}_u$ particles are injected into the upstream per one electron (or positron) heated in the shock. They join the upstream, cool down, and then go through the shock like the primary electron did, closing the cycle. This cycle allows the upstream $e^\pm$ population to reproduce itself with the amplification factor $\mathcal{M}_u > 1$.

The breeding of upstream pairs, however, is limited, because $\gamma_{\text{th,e}} \propto Z_\pm^{-1}$. The increasing upstream $Z_\pm$ eventually reduces $\gamma_{\text{th,e}}$ to $\gamma_{\text{cr}}$, which is marginally sufficient for production of IC photons capable of converting ahead of the shock. Then a self-consistent situation is achieved with $\mathcal{M}_u \approx 1$—the upstream pair population is replenished with no further growth. The pair loading factor ahead of the shock is then given by

$$Z_\pm \approx f_e (\gamma_0 - 1) \frac{m_p}{m_e} \sim 10^2 f_e \frac{\gamma_{\text{cr}}}{20} \left(\frac{\gamma_0}{\gamma_0 - 1}\right).$$  \hfill (74)

The critical value of $\gamma_{\text{th,e}}$ can be estimated for a typical spectrum of radiation carried by GRB jets, which peaks around 10 keV in the fluid frame. Then $\gamma_{\text{cr}} \sim 20$ is capable of producing a large number of IC photons with $E_{\text{IC}} > 3$ MeV.

The reduction of $\gamma_{\text{th,e}}$ to $\gamma_{\text{cr}}$ implies that the downstream $e^\pm$ cascade has only one generation and a modest multiplicity $\mathcal{M}_\pm \sim 3-5$. The downstream pair loading factor is much higher, $Z_\pm \gg \mathcal{M}_\pm$, due to the breeding of pairs ahead of the shock.

In an RMS, with or without a collisionless subshock, the upstream is decelerated by scattering the radiation flux from the downstream. The absorption of MeV photons implies an additional deceleration effect. It involves the upstream deposition of momentum $\sim Z_\pm m_e c$ per ion, which may approach $\sim 0.1$ of the ion momentum and is insufficient to significantly change the velocity profile of the shock wave.

This situation is in contrast with the previously studied $e^\pm$ creation ahead of external blast waves from GRB explosions (Thompson & Madau 2000; Beloborodov 2002). In that case, the shock is driven by ultrarelativistic ejecta into a low-density ambient medium of rest mass $mc^2 \sim E_{\text{IC}}/Y^2$, where $E_{\text{IC}}$ is the explosion energy. The average GRB photon propagating ahead of the shock has energy comparable to 1 MeV in the rest frame of the ambient medium. Then the impact of pair creation and radiation pressure on the external medium is huge. The MeV radiation front is capable of sweeping up the external medium and even clearing a vacuum cavity ahead of the ejecta (Beloborodov 2002). Such effects cannot occur in internal shocks because they are much weaker and involve a much lower contrast in density.

### 6.3. Energy Runaway?

Pair breeding described above involves electrons and photons of MeV energies and does not generate very high-
energy particles. The most energetic particles have Lorentz factors $\gamma_i \sim \gamma_0$, immediately behind the collisionless shock, and $e^\pm$ pairs injected into the upstream have mildly relativistic $\gamma_i \sim 2$. Before reaching the collisionless shock, these pairs lose energy through Coulomb collisions with the background plasma and IC cooling (synchrotron cooling of mildly relativistic electrons is suppressed by self-absorption). Compton cooling results from scattering of background photons of energies $E_\gamma < 100$ keV, and hence the upstream produces IC photons with energies $E_{IR} \sim \gamma_0^2 E_\gamma < m_e c^2$. Thus, MeV radiation is only generated immediately behind the collisionless shock, with no amplification cycle in photon energy.

The situation could change, however, if the shock is ultrarelativistic, $\gamma_0 \gg 1$. Then pair breeding occurs with a high energy gain per cycle, tapping into the kinetic energy of upstream motion relative to downstream. This can lead to a runaway cycle of photon energy boosting, which was noticed and described as a “converter mechanism” by Derishev et al. (2003) and “electromagnetic catastrophe” by Stern (2003). The cycle involves IC emission followed by photon propagation across the collisionless shock, $e^\pm$ creation on the other side, and IC emission from the created pair back toward the downstream. This effectively implies an exchange of IC photons between the upstream and downstream accompanied by the energy boost $\sim \gamma_0^2$. The runaway occurs if the cycle is closed with the mean energy for the final photon energy exceeding the initial photon energy.

Consider an initial IC photon of energy $\epsilon_i m_e c^2$ in the downstream frame, and suppose the photon overtakes the shock and converts to an $e^\pm$ pair. The created particles have Lorentz factors $\gamma_i \sim \gamma_0 \epsilon_i/2$ in the upstream frame, and cool in the Klein–Nishina regime if $\gamma_\gamma > 1$ (where $\gamma_\gamma \epsilon_i c^2 \sim 10$ keV describes the peak of the radiation spectrum in the fluid frame). Then the new IC photons emitted by the $e^\pm$ pair and viewed from the downstream frame have energies $\epsilon_j \sim \gamma_0 \gamma_j/2 \sim (\gamma_0^3/4) \epsilon_i$. The cycle is completed when the photon $\epsilon_j$ propagates into the downstream, creates a pair with $\gamma_j \sim \epsilon_j/2$, and this pair produces IC photons with $\epsilon_f \sim \gamma_j/2 \sim (\gamma_0^3/4) \epsilon_j$. This rough estimate suggests that a runaway cycle, $\epsilon_f > \epsilon_i$, requires $\gamma_0 > 4$.

It also requires two other conditions:

1. The electron cooling free path $l_{IC}$ should not exceed the free path of photons it produces, $L_{\gamma \gamma}$; otherwise the IC photon is absorbed before it has a chance to cross the shock. This condition is not satisfied for high-energy IC photons, $\epsilon \gg \epsilon_p^{-1}$, if the target radiation spectrum at $\epsilon_i < \epsilon_p$ has the photon index $\alpha = d \ln n_\gamma/d \ln \epsilon_i < 0$. If $\alpha > 0$, IC scattering and $\gamma\gamma$ absorption of photons with $\epsilon > \epsilon_p^{-1}$ involve the same main targets $\epsilon_i \sim \epsilon_p$ with comparable cross sections $\sigma_{\gamma\gamma} \sim \sigma_{IC} \sim (\epsilon_p)^{-1} \sigma_{\gamma\gamma}$; then $l_{IC} \sim L_{\gamma \gamma}$.

2. The magnetic fields should be weak; otherwise synchrotron losses become increasingly dominant over IC emission in the deep Klein–Nishina regime and suppress the production of high-energy IC photons.

In summary, the runaway of photons with growing energies requires an ultrarelativistic collisionless shock and weak magnetic fields. This combination is unlikely at large optical depths where the formation of strong collisionless shocks requires strong magnetic fields.

6.4. Ion Cooling

The energy fraction kept by the ions behind the collisionless shock, $1 - f_\gamma$, is not easily radiated—an ion cannot directly radiate its energy, because of its large mass $m_i$. The ions gradually lose their energy through Coulomb collisions with $e^\pm$ or through nuclear collisions. Both processes are relatively slow, and the ion cooling can be a bottleneck for postshock conversion of heat to radiation.

The $e^\pm$ plasma behind the shock quickly becomes much colder than the ions. Frequent Compton scattering enforces kinetic equilibrium of electrons (and positrons) with local radiation at the Compton temperature $k T_{IC} \sim 10–50$ keV, which corresponds to a thermal speed around $0.3c$. In the first approximation, the hot, mildly relativistic ions behind the collisionless shock view electrons as targets at rest, and the ion cooling timescale due to Coulomb collisions is approximately given by (e.g., Ginzburg & Syrovatskii 1964)

$$t_{ICool} = \frac{(\gamma_{th} - 1) m_i c^2}{E_{ICool}} = \frac{2 \beta_{th}(\gamma_{th} - 1) m_i}{3 \ln \Lambda \sigma_{in} n_\pm m_e c},$$

where $n_\pm$ is the local density of electrons and positrons and $\Lambda \approx 20$ is the Coulomb logarithm.

During time $t_{ICool}$, the hot ions are advected through the distance

$$l_{ICool} = v_i t_{ICool} = \frac{v_0}{3} t_{ICool},$$

where $v_i \approx v_0/3$ is the velocity of the downstream relative to the collisionless jump. The Thomson optical depth of the ion cooling region is

$$\tau_T = \sigma_T n_\pm l_{ICool} = \frac{2 m_i \beta_{th} \beta_{th}(\gamma_{th} - 1)}{9 m_e \ln \Lambda}.$$  

A significant fraction of radiation produced by the ion cooling at distance $l_{ICool}$ behind the shock can diffuse back into the upstream if $\tau_T v_T/c \lesssim 1$. Otherwise, radiation is trapped and advected away from the shock, missing the chance to affect the upstream velocity profile. We conclude that the ion heat produced by the collisionless jump is effectively lost for the RMS if

$$\frac{2 m_i \beta_{th} \beta_{th}(\gamma_{th} - 1)}{27 m_e \ln \Lambda} > 1.$$

This condition is satisfied for collisionless jumps with amplitude $p_0 = \gamma_0 \beta_0 \gtrsim 1$. In this case, the delay in ion cooling tends to reduce the radiative precursor and increase the amplitude of the collisionless jump.

The mildly relativistic ion (proton) temperature behind a strong subshock, $k T \sim 1$ GeV, leads to inelastic nuclear collisions between the protons, producing pions. Half of the pion energy is lost to neutrino emission while the other half is converted to $e^\pm$ with Lorentz factors $\gamma_\gamma \approx m_p/m_e \sim 300$. This injection of relativistic pairs by $p-p$ collisions sustains an $e^\pm$ cascade in the downstream region of thickness $l_{COOL}$. The cascade is similar to that triggered by inelastic n–p or n–n collisions in a neutron-loaded jet (Derishev et al. 1999; Beloborodov 2010).

The timescale for inelastic $n-p$ collisions is $t_{np} \approx (\sigma_{np} \sigma_{n-p})^{-1}$, where $\sigma_{n-p} = f_{n-p} \sigma_n$ is a substantial fraction of the nuclear cross section $\sigma_n \approx \sigma_T/20$. The ion heat lost to inelastic collisions before it is given to $e^\pm$ via Coulomb collisions is determined by
the ratio
\[
\frac{I_{\text{coll}}}{I_{pp}} = \frac{2\beta_{\text{th}}(\gamma_{\text{th}} - 1) m_e \sigma_{\text{coll}}}{3 \ln A Z_{\pm} m_e \sigma_T} \sim \frac{1}{Z_{\pm}}.
\] (79)

One can see that \(e^\pm\) loading reduces the role of p-p collisions behind the shock.

7. Transformation of Shocks near the Photosphere

As an RMS emerges from the photosphere of the outflow, it must transform into a pure collisionless shock: radiation becomes decoupled from the plasma and the shock must be sustained by collective (collisionless) plasma processes. This transformation occurs through the growth of the collisionless subshock inside the RMS.

A key feature described in Sections 5.2 and 6 is that the shocks are dressed in \(e^\pm\) pairs. This delays their transition to transparency and thus delays the transition to a pure collisionless state. While the upstream and far downstream regions are already transparent to radiation, the shock itself remains opaque until its pair dress becomes optically thin. The estimated pair loading factor \(Z_{\pm}\) inside the shock front is comparable to \(10^2\), and hence the effective photospheric radius should be increased by a factor \(\sim 10^5\). The shock “carries” the photosphere with it and keeps radiating right at the photosphere rather than crosses it.

Pair creation is strong even before the shock approaches the photosphere. A weakly magnetized relativistic shock generates pairs through bulk Comptonization, as described in Section 5.2. In the magnetized case, pairs are generated by the collisionless subshock, as described in Section 6. In any case, when the shock attempts to emerge from the photosphere and become a pure collisionless jump, pair creation with a large \(Z_{\pm}\) is inevitable.

Details of the photospheric transition depend on the shock amplitude \(p_0 = \gamma_0 \beta_0\) and magnetization \(\gamma\). The radius \(R_{\pm}\) where the shock eventually becomes transparent satisfies the condition \(R_{\pm} \sim Z_{\pm} R_e\), where \(R_e\) is the radius of the photosphere in the absence of pair creation. Note that our estimates for \(Z_{\pm}\) assumed that most MeV photons (\(\sim \Gamma\) MeV in the lab frame) emitted by the shock at a radius \(r\) convert to pairs with a free path \(\ll r\) in the lab frame, which corresponds to \(E_{\gamma} \ll r/\Gamma\) in the fluid frame. Taking into account that \(\sigma_{\gamma \gamma} \sim 0.1 \sigma_T\), this requires a sufficiently large “compactness” parameter,

\[
l = \frac{U_{\text{rad}}}{m_e c^2} \sigma_T \frac{r}{\Gamma} \gg 10,
\] (80)

where \(U_{\text{rad}} = L_{\text{rad}} / 4\pi r^2 c\Gamma^2\) is the radiation energy density in the fluid frame and \(L_{\text{rad}}\) is the isotropic equivalent of the observed GRB luminosity. The compactness parameter is related to the characteristic Thomson optical depth of the outflow \(\tau_T = n_e \sigma_T r/\Gamma\),

\[
l \sim \frac{n_e}{Z_{\pm}} \frac{\bar{E}}{\Gamma m_e c^2} \sigma_T \sim \frac{10^3}{Z_{\pm}} \tau_T,
\] (81)

where \(\bar{E} \sim 1\) MeV is the average photon energy measured in the static lab frame, and \(n_e / n \sim 10^5\) is the typical photon-to-baryon ratio (we use the numerical values typical for GRBs). If \(l \lesssim 10^{-30}\), the \(\Gamma\) MeV photons do not convert to pairs; only photons of significantly higher energies are quickly absorbed. A rough estimate for \(R_{\pm}\) may be obtained by combining Equations (74) and (81), which gives \(R_{\pm} \sim 30 R_e\). Its exact value depends on the parameters of the explosion.

A moderate magnetization of the flow \(\gamma \sim 0.1\) implies that the collisionless shock expanding from \(R_e\) to \(R_{\pm}\) is a strong source of synchrotron photons. Unlike the IC photons (many of which convert to pairs) the synchrotron photons are soft and dominate the low-energy part of the shock radiation spectrum. Overall, the photospheric radiation released by the \(e^\pm\)-dressed shock should have a broad nonthermal spectrum around \(E \sim 1\) MeV in the lab frame. It extends from the synchrotron self-absorption energy (well below 1 MeV) to the GeV break shaped by \(\gamma \gamma\) absorption.

Detailed calculations of the emitted spectrum are deferred to a future paper. Here we note that the structure and observational appearance of the shock emerging from \(R_e\) are affected by the development of radiation anisotropy in the fluid frame. In any relativistic outflow, radiation develops a strong forward beaming in the fluid frame at optical depths \(\tau_T \lesssim 10\) (Beloborodov 2011). As the \(e^\pm\)-dressed shock expands from \(R_e\) to \(R_{\pm}\), its radiation maintains a strong beaming. This effect breaks the symmetry between shocks propagating forward and backward relative to the plasma outflow (shocks form in pairs propagating in opposite directions relative to the fluid, see Section 2). The downstream of a backward internal shock is radially ahead of the upstream; hence the forward beaming reduces the efficiency of radiation diffusion from the downstream to the upstream. This accelerates the development of the collisionless shock and also influences the effective optical depth \(\tau_T\) of the \(e^\pm\) dress. In contrast, for a forward-propagating shock the downstream is radially behind the upstream. Then beaming assists in sending photons into the upstream with a decreasing angle relative to the radial direction.

8. Nuclear Collisional Dissipation

8.1. Neutral Particles in Sub-photospheric Shocks

In general, shock formation occurs when the steepening of the compressive wave is stopped by momentum (or heat) transfer due to the finite mean free path of particles. In a multi-component fluid, the importance of different components for the shock structure is determined by their contributions to viscosity and thermal conductivity. Both are controlled by the diffusion coefficient,

\[
D = \frac{1}{3} \ell \bar{v},
\] (82)

where \(\ell\) is the mean free path and \(\bar{v}\) is the characteristic thermal speed of a given species of particles. In particular, the viscosity coefficient created by each species may be estimated as

\[
\mu = D \xi \frac{H}{c^2},
\] (83)

where \(\xi\) is the fractional contribution of the species to the relativistic enthalpy density of the flow \(H\), which includes the rest-mass energy (\(H/c^2\) serves as the effective inertial mass density in relativistic hydrodynamics).

Consider now shock formation deep below the photosphere. A simple comparison of viscosity coefficients of different components of the preshock flow allows one to judge their importance for the shock structure. In a first approximation, electrons, positrons, ions, and magnetic fields move together as a single, strongly coupled fluid. In contrast, neutral particles—
photons and especially neutrons—have large free paths and a large diffusion coefficient, which leads to significant transfer of momentum and energy between the upstream and downstream, shaping the profile of the shock wave.

The consideration of diffusion coefficients assumes that the particles are not decoupled from the flow, i.e., their mean free-path time $\ell / \nu$ is smaller than the age of the flow (measured in its rest frame). Neutrinos may carry a significant fraction of the flow energy in GRBs, but they escape freely and do not contribute to viscosity. Photons contribute to viscosity at radii $r \lesssim R_s$, and neutrons at radii $r \lesssim R_n$ where $R_n \approx R_s (\sigma_\nu / Z_\nu \sigma_T)$ is the neutron decoupling radius.

8.2. Neutron-mediated Shock Wave

GRB jets are expected to carry a significant number of free neutrons (Derishev et al. 1999; Beloborodov 2003). Their $\beta$-decay is delayed proportionally to the outflow Lorentz factor $\Gamma$ and occurs at a characteristic radius $R_\beta \approx 8 \times 10^{13} (\Gamma / 300)$ cm, which is much larger than the typical photospheric radius $R_\nu \sim 10^{12} \sim 10^{13}$ cm. Neutrons are coupled to the plasma through nuclear collisions with cross section $\sigma_\nu \approx 3 \times 10^{-26}$ cm$^2$, which is 20 times smaller than the Thomson cross section. Neutrons have the largest mean free path, carry a significant fraction of the flow momentum, and hence can dominate the flow viscosity, affecting the formation of shocks. Neutron migration across the shock assists the momentum exchange between the upstream and downstream on a scale comparable to the neutron mean free path $\ell_n$, shaping a broad, “neutron-mediated” shock wave.

This wave can have a strong subshock mediated by radiation, because the neutron collisions alone are unable to stop the velocity profile everywhere from steepening. If the neutron fraction of the outflow momentum is modest, the main momentum jump in the wave occurs in the subshock. Then the wave is better described as an RMS with a neutron precursor rather than a neutron-mediated shock. The RMS itself can have a strong collisionless subshock, depending on the flow magnetization. The resulting wave structure is shown schematically in Figure 11.

8.3. Nuclear Collisions Around an RMS

Let us first consider small radii where the outflow rest-mass energy is dominated by radiation, $U_{\text{rad}} \gg \rho c^2$. Then the energy of neutrons (including their rest mass) is small compared with the energy dissipated in the RMS. The RMS thickness $\ell_\text{th}$ is comparable to the photon mean free path $\ell_\text{ph}$. It is much smaller than the neutron mean free path $\ell_n$.

$$\frac{\ell_n}{\ell_\text{ph}} \sim Z_\perp \sigma_T / \sigma_\nu \approx 20 Z_\perp,$$

where $Z_\perp = n_\perp / n \geq 1$ determines the reduction in the photon free path in the plasma enriched by $e^\pm$ pairs.

Neutrons brought by the upstream view the RMS as a discontinuity in the fluid velocity profile. They cross it ballistically and dissipate their energy $(\gamma_0 - 1) m_e c^2$ and momentum $p_0 m_e c$ at the characteristic distance $\sim \ell_n$ downstream of the shock. Their collisions with the downstream nuclear matter create a relativistically “hot” neutron (and ion) component embedded in the photon gas. Some of the hot neutrons propagate back into the upstream and collide there. They form a precursor of the shock—a “neutron pillow” that somewhat decelerates the upstream and thus reduces the strength of the RMS.

As long as the neutron rest-mass density $\rho_n c^2$ is small compared with the enthalpy of the upstream photon gas, $4 P_\text{th}$, the effect of neutron transport on the RMS amplitude is small. For example, consider a radiation-dominated jet with an asymptotic (saturation) Lorentz factor $\Gamma_\text{sat}$ at a small radius where the Lorentz factor is $\Gamma = 0.2 \Gamma_\text{sat}$. At this early stage, 80% of the jet energy is carried by radiation, and about 20% is carried by baryons (neglecting here the energy of the magnetic field and $e^\pm$). Assuming that half of the baryons are neutrons, one can estimate $\rho_n c^2 / P_0 \approx 0.1$, so roughly 10% of the shock energy is dissipated by nuclear collisions in this example. This dissipation occurs in the region of thickness $\sim \ell_n$ around the RMS. Even the modest amount of collisional dissipation at small radii (large optical depths) is important because it efficiently generates photons, as discussed below.

As the expanding and accelerating outflow experiences adiabatic cooling, the ratio $\rho_n c^2 / P_0$ increases and the role of collisional dissipation grows. When the Lorentz factor $\Gamma$ approaches its maximum value $\Gamma_\text{sat}$, i.e., when the radiation density $U_{\text{rad}}$ becomes comparable to $\rho c^2$, a large fraction of the shock energy becomes dissipated through nuclear collisions. The wave front of thickness $\sim \ell_n$—the neutron-mediated shock wave—propagates as long as the flow is opaque to nuclear collisions. When the wave approaches the neutron decoupling radius $R_n$, the shock energy becomes dissipated through nuclear collisions.

The RMS itself is not capable of producing ultrarelativistic particles, and so collisional dissipation around the RMS plays a
key role in this respect. Even in moderately relativistic shocks the neutron collisions are energetic enough to be inelastic. Such collisions produce mildly relativistic pions, which quickly decay into ultrarelativistic $e^\pm$ with Lorentz factors $\gamma_e \sim m_e/m_e$ (Derishev et al. 1999). This generates an IC cascade that produces copious $e^\pm$ pairs. The resulting pair loading factor $Z_\pm = n_\gamma/n$ is comparable to 10 as long as the jet magnetization parameter is below $0.1$ (Beloborodov 2010; Vurm et al. 2011). Far downstream of the shock, where the baryons have cooled, the inelastic collisions and the cascade end; here the pairs annihilate, if they still have time to do so before freezing out. The freeze-out happens when the outflow expansion timescale becomes shorter than the annihilation timescale; this occurs when the flow approaches the photosphere (Beloborodov 2010).

Note that neutrons can convert to protons (and protons to neutrons) in inelastic nuclear collisions. This enables the “converter” mechanism for baryon acceleration proposed by Derishev et al. (2003). Numerical results of Kashiya et al. (2013) suggest that the converter mechanism becomes efficient for ultrarelativistic shocks, $\gamma_0 \gtrsim 4$.

9. Discussion

This paper focused on sub-photospheric internal shocks in relativistic explosions, their dissipation mechanism, and their structure. One question of observational interest is whether the shocks are capable of producing ultrarelativistic electrons. Energetic electrons at large optical depths emit synchrotron radiation (without self-absorption) and thus boost the photon number carried by the flow, which is later released at the photosphere. Another interesting question is how the shock evolves and radiates as it approaches the photosphere.

As long as the flow is opaque and radiation dominates its energy density, photon transport plays a leading role in shaping the shock front. Its thickness is then comparable to the photon mean free path. The RMS is not capable of electron acceleration by the standard Fermi mechanism, since the electron radiates its energy faster than it can cross the shock. Photons experience significant energy gains by crossing the shock back and forth multiple times. This “bulk Comptonization” upscatters photons up to the MeV band (in the fluid frame); further upscattering is hindered by the energy loss due to electron recoil in scattering. The photon upscattering beyond $\sim 1$ MeV is also stopped by the absorption reaction $\gamma + \gamma \rightarrow e^- + e^+$. The mildly relativistic $e^\pm$ pairs produced immediately cool down due to fast Coulomb and IC losses.

All this would suggest that sub-photospheric shocks are inefficient in producing particles with energies $E \gg m_e c^2$ in the fluid frame. However, a more realistic shock picture differs significantly from the simple RMS, in particular when one takes into account that the outflow carries magnetic fields and free neutrons (see Figure 11). The shock wave is capable of generating ultrarelativistic electrons in two ways.

1. A strong collisionless subshock forms in the RMS. We have shown that this is inevitable (even deep below the photosphere) if the flow is sufficiently magnetized. A mildly relativistic collisionless subshock heats the electrons to an ultrarelativistic temperature $T_e$. Their IC emission breeds $e^\pm$ pairs in the upstream and regulates the shock structure to a self-consistent state with the postshock temperature $kT_e \sim 10m_e c^2$ (Section 6.2). This temperature is high enough to generate interesting synchrotron radiation without self-absorption, which is strong for electrons with Lorentz factors $\gamma_e \lesssim 10$ (Vurm et al. 2011). The subshock fails to generate particles with $\gamma_e > 10$ if it is weak (the flow is weakly magnetized), while pair loading remains strong, $Z_\pm \sim 10^2$, due to the conversion of bulk-Comptonized MeV photons (Section 5.2).

2. Inelastic nuclear collisions inject $e^\pm$ pairs with Lorentz factors $\sim m_n/m_e \sim 300$ in the fluid frame. This mechanism becomes particularly efficient if the outflow carries free neutrons, because they can migrate across the RMS, making the shock wave partially mediated by neutrons (Figure 11).

The synchrotron losses of energetic electrons generated by either mechanism imply significant photon production. The synchrotron photons may carry a small fraction of radiation energy, but their number is significant. GRB jets tend to experience photon starvation in the “Wien zone” at optical depths $10^2 < \tau_\gamma < 10^5$ (Beloborodov 2013). In this zone, the heated photons have a Wien (rather than Planck) spectrum, i.e., a Bose–Einstein distribution with a nonzero chemical potential. The production of low-energy photons is followed by their quick Comptonization to the Wien peak. The addition of photons shifts the peak to lower energies, as the energy per photon is reduced. This effect regulates the observed peak position of the GRB spectrum that is eventually released at the photosphere (Vurm & Beloborodov 2016). Synchrotron photons produced outside the Wien zone, i.e., at smaller optical depths $\tau_\gamma < 10^2$, are only partially Comptonized toward the Wien peak and form the low-energy part of the prompt GRB spectrum with the characteristic photon index $\alpha \sim 1$ (see Thompson & Gill 2014; Vurm & Beloborodov 2016).

Nuclear collisions in the shock front also generate neutrinos. Neutrino emission from migrating and colliding neutrons in GRB jets was previously discussed in some detail (Derishev et al. 1999; Balck & Mészáros 2000; Mészáros & Rees 2000a). The typical energy of neutrinos produced by this mechanism is $\sim 10m_\nu c^2 \gtrsim 10$ GeV and they are detectable by IceCube (Bartos et al. 2013; Murase et al. 2013). Sub-photospheric internal shocks were also proposed to emit ultrahigh-energy neutrinos (Mészáros & Waxman 2001; Razzaque et al. 2003). This proposal assumed efficient ion acceleration by the Fermi diffusive mechanism. This mechanism does not operate in an RMS. Diffusive acceleration is also suppressed in the collisionless subshock with a transverse magnetic field, which advects the particles downstream before they have a chance to cross the shock many times (Sironi & Spitkovsky 2011). An oblique magnetic field could help this process to occur.

Detailed studies of sub-photospheric shock structure require numerical simulations. Previous work focused on the search for a steady-state solution for the RMS (e.g., Levinson & Bromberg 2008; Budnik et al. 2010; T0lsto et al. 2015), which may be found by iteration. Instead, we suggest two techniques that permit direct time-dependent simulations of shock formation, as demonstrated in Sections 4 and 5. Our simulations are set up to follow the evolution of an internal compressive wave, which leads to formation of a pair of shocks and their subsequent quasi-steady propagation. The shock structure is obtained from first principles, by simulating the time-dependent radiative transfer in the moving plasma.
We have implemented two methods for such simulations: (1) solving the radiative transfer equation coupled to the flow dynamics, and (2) tracing individual photons and their interaction with the moving plasma using the Monte-Carlo technique. Both methods show the structure of the established shock wave (Figures 8 and 9) and reproduce the jump conditions described in Section 3. The simulations verified the formation of a strong subshock in a magnetized RMS. In particular, in GRB jets, a moderate magnetization \( \sim 0.1 \) is sufficient to generate a strong subshock.

A curious feature indicated in Figure 11 and discussed in Section 6 is the delayed cooling of the ions heated in the collisionless subshock. If the subshock is relativistic, the ion cooling length \( l_{\text{cool}} \) can exceed the RMS thickness. When the ion energy is finally radiated, the radiation produced is trapped and advected downstream, missing the chance to diffuse upstream and affect the shock velocity profile. The delayed ion cooling also implies that some ions experience inelastic nuclear collisions and emit neutrinos even in the absence of a free neutron component.

We have estimated the pair loading factor \( Z_{\pm} \sim 10^2 \) in the RMS with or without a collisionless subshock. The \( e^\pm \) pairs are produced in collisions between MeV photons, which are generated by two mechanisms. RMS without a strong collisionless subshock produces MeV photons only through bulk Comptonization in a relatively cold converging flow, which is close to Compton equilibrium with local radiation at \( KT \ll m_e c^2 \). In the presence of a collisionless jump, MeV photons are produced by IC cooling of \( e^\pm \) heated in the jump to \( KT \sim 10m_e c^2 \).

As the GRB jet expands from the central engine the structure of internal shocks and the dissipation mechanism change. As long as the jet energy is dominated by radiation, \( U_{\text{rad}} \gg \rho c^2 \) and \( U_{\text{rad}} \gg U_{\text{fr}} \), the shock structure is described by a unique solution, which has no collisionless subshock (Figure 8). Bulk Comptonization in such deep sub-photospheric shocks is also suppressed, and their structure is conveniently described by the “force-free” radiative transfer solution for the bolometric intensity. This regime can occur at small radii in GRB jets, where the outflow Lorentz factor \( \Gamma \ll \Gamma_{\text{sat}} \).

Nuclear dissipation due to neutron migration across the RMS increases with \( \Gamma \) and approaches its maximum near the radius of Lorentz factor saturation. It remains high until the neutron decoupling radius \( R_n \), and then it declines as \( (r/R_n)^{-1} \).

The shock “breakout” at the photosphere occurs through the growth of the collisionless subshock in the RMS until radiation completely decouples from the plasma. Eventually the entire velocity jump becomes mediated by collective plasma processes, regardless of magnetization. This somewhat resembles the shock breakout in nonrelativistic supernova explosions (Waxman & Loeb 2001; Giacinti & Bell 2015). However, there is a special feature: \( e^\pm \) pair creation near the collisionless shock sustains an optical depth \( \tau_r \gtrsim 1 \) even after the background electron–ion plasma becomes transparent. Effectively, the shock carries the photosphere with it until it expands by an additional factor \( \sim 30 \), continually producing photospheric emission. This emission will be observed as a prominent pulse of nonthermal radiation in the GRB light curve.

In the observer frame, the emission from the \( e^\pm \)-dressed shock extends up to the GeV band, where \( \gamma-\gamma \) absorption shapes a break in the spectrum. The high-energy photospheric pulses overlap, in observer time, with the GeV flash that is produced by the \( e^\pm \)-loaded external blast wave at a larger radius \( R \sim 10^{16} \text{ cm} \) (Beloborodov et al. 2014; Hascoët et al. 2015). These pulses may explain the observed variable component of GeV emission superimposed on the smooth GeV flash at early times (Ackermann et al. 2013).

Internal shocks are a particularly promising heating mechanism for jets that are not magnetically dominated. Comparison of detailed models of photospheric radiation with observed GRB spectra suggests a moderate magnetization in the sub-photospheric region, \( \sigma \sim 10^{-2} - 10^{-1} \) (Vurm & Beloborodov 2016). However, this does not exclude a stronger magnetization close to the central engine, allowing for magnetic dissipation that reduces \( \sigma \) as the jet expands. In this scenario, the early heating would be dominated by magnetic dissipation. Even in this regime shocks can occur and dissipate significant energy, as follows from the jump conditions discussed in Section 3.

I would like to thank Hirotaka Ito, Christoffer Lundman, and Indrek Vurm for discussions and comments on the manuscript. This work was supported by NSF grant AST-1412485, NASA grant NNX15AE26G, and a grant from the Simons Foundation (###446228, Andrei Beloborodov).

### Appendix

#### Location of Shock Formation in a Supersonic Compressive Wave

A supersonic wave converging toward the caustic \( x = 0 \) forms a pair of shocks at the Lagrangian coordinates \( \pm x_0^\ast \). The value of \( x_0^\ast \) can be estimated as follows.

Consider the streamlines that experience smooth compressive deceleration, due to the conversion of kinetic energy to enthalpy. For a streamline with a given Lagrangian coordinate \( x_0 \) the characteristic time \( t_{\text{dec}}(x_0) \) and location \( x_{\text{dec}}(x_0) \) where deceleration occurs are given by Equations (9) and (10). Next, note that the smooth compressive deceleration to a subsonic speed is possible as long as the density of the gas accumulated in the subsonic region, \( \rho_{\text{dec}} \), is comparable to the density of the ballistic flow approaching it,

\[
\rho_{\text{bal}} \approx \frac{\rho_0}{1 + \rho_0 t_{\text{dec}}}.
\]

Since \( \rho \propto \rho_0 \) and pressure in the subsonic region is not far from uniform (it tends to equilibrate on the sound crossing time), \( \rho_{\text{dec}} \) is roughly uniform,

\[
\rho_{\text{dec}} \approx \frac{x_0}{x_{\text{dec}}} \rho_0. \quad (85)
\]

The density ratio is given by

\[
f(x_0) \equiv \frac{\rho_{\text{bal}}}{\rho_{\text{dec}}} \approx 1 + \rho_0 t_{\text{dec}}/x_0. \quad (86)
\]

Using Equation (9), one can exclude \( t_{\text{dec}} \) and obtain

\[
f(x_0) \approx 1 - \left( \frac{\rho_0}{\rho_0 x_0} - 1 \right) \left[ \frac{(\alpha - 1)\rho_0^2}{2c_0^2} \right]^{1/(\alpha - 1)} - 1 \]. \quad (87)

The unity in the last term (in square brackets) may be neglected for streamlines with \( \rho_0 \gg c_0^2 \). One can see that \( f < 1 \) and \( f \) is close to unity for small \( |x_0| \). It drops sharply when \( |x_0| \) exceeds some characteristic \( x_0^\ast \), and formally even changes sign. The characteristic \( x_0^\ast \) may be estimated from the condition \( f \sim 0 \).
For waves with amplitudes $v_{\text{max}} \gg c_0$ one finds that $x_0^* \approx 0$ is much smaller than the wavelength. Therefore, in the calculation of $x_0^*$ one can use the Taylor expansion of the velocity profile around $x_0 = 0$,

$$v_0(x_0) = -a x_0 + \frac{b}{6} x_0^3 + O(x_0^4),$$

where $a = v_0'(0)$ and $b = v_0''(0)$. Here we took into account that the second derivative $v_0''$ vanishes at $x_0 = 0$ (recall that we chose the zero $x$-coordinate at the location of the caustic where $-v_0'$ is maximum, see Section 2.1). A linear expansion $v_0 = -ax_0 + O(x_0^2)$ would not be sufficient—a uniform $v_0(x_0)$ would imply a uniform compression of the ballistic flow, with no pressure gradient that could cause deceleration. One can also see from Equation (87) that it is the deviation from the linear velocity profile $v_0(x_0) \approx 1 \approx b x_0^3/3a$ that controls the drop of $f$ below unity, and hence controls shock formation.

For the supersonic streamlines with $v_0^2 \gg c_0^2$ Equation (87) yields the relation

$$1 - f \approx \frac{b}{3a} x_0^3 \left( \frac{1}{(2\pi)^{1/2}} \right)^{1/(2\pi - 1)},$$

and hence

$$x_0^* \approx \frac{3^{\alpha - 1}}{\alpha - 1} 2^{\gamma - 3} b^{1 - \alpha}.$$

The coefficients $a$ and $b$ in the Taylor expansion (88) can be estimated as $a \approx p_{\text{max}} c/L$ and $b \approx p_{\text{max}} c/L^3$ for a smooth initial profile of the wave; these relations are exact for a sine profile $p_0(x_0) = -p_{\text{max}} \sin(x_0/L)$. Substitution into Equation (90) gives

$$\frac{x_0^*}{L} \approx \chi \left( \frac{c_0}{c_{\text{max}}} \right)^{1/\alpha},$$

where $\chi \approx 3^{(\alpha - 1)/2\alpha} [2/(\alpha - 1)]^{1/2\alpha} \sim 2$ for the relevant range of $4/3 < \alpha < 2$. Numerical simulations (similar to the sample model shown in Figure 3, with different $\alpha$ and $c_0$) provide the accurate location of shock formation and confirm the scaling predicted by Equation (91), with a slightly larger numerical coefficient $\chi \approx 3.4$.

References

Ackermann, M., Ajello, M., Asano, K., et al. 2013, ApJS, 209, 11
Amato, E., & Arons, J. 2006, ApJ, 653, 325
Bahcall, J. N., & Mészáros, P. 2000, PhRvL, 85, 1362
Bartos, I., Beloborodov, A., Hurley, K., & Marka, S. 2013, PhRvL, 110, 241101
Beloborodov, A. M. 1999, MNRAS, 305, 181
Beloborodov, A. M. 2002, ApJ, 565, 808
Beloborodov, A. M. 2003, ApJ, 588, 931
Beloborodov, A. M. 2010, MNRAS, 407, 1033
Beloborodov, A. M. 2011, ApJ, 737, 68
Beloborodov, A. M. 2013, ApJ, 764, 157
Beloborodov, A. M., Hascoët, R., & Vurm, I. 2014, ApJ, 788, 36
Beloborodov, A. M., & Uhm, Z. L. 2006, ApJ, 651, L1
Blandford, R. D., & Payne, D. G. 1981, MNRAS, 194, 1041
Bromberg, O., Mikolaitis, Z., & Levinson, A. 2011, ApJ, 733, 85
Budnik, R., Katz, B., Sagiv, A., & Waxman, E. 2010, ApJ, 725, 63
Daigne, F., & Mochkovitch, R. 1998, MNRAS, 296, 275
de Hoffmann, F., & Teller, E. 1950, PhRv, 80, 692
Derishev, E. V., Aharonian, F. A., & Kocharovsky, V. V. 2003, PhRvD, 68, 043003
Derishev, E. V., Kocharovsky, V. V., & Kocharovsky, V. V. 1999, ApJ, 521, 640
Giacinti, G., & Bell, A. R. 2015, MNRAS, 449, 3693
Ginzburg, V. L., & Syrovatskii, S. I. 1964, The Origin of Cosmic Rays (Oxford: Pergamon)
Hascoët, R., Vurm, I., & Beloborodov, A. M. 2015, ApJ, 813, 63
Hoshino, M., Arons, J., Gallant, Y. A., & Langdon, A. B. 1992, ApJ, 390, 454
Ito, H., Matsumoto, J., Nagataki, S., Warren, D. C., & Barkov, M. V. 2015, ApJ, 814, L29
Kashiyama, K., Murase, K., & Mészáros, P. 2013, PhRvL, 111, 131103
Kobayashi, S., Piran, T., & Sari, R. 1997, ApJ, 490, 92
Landau, L. D., & Lifshitz, E. M. 1959, Fluid Mechanics (Oxford: Pergamon)
Lazzati, D., Mori, B. J., Margutti, R., & Begelman, M. C. 2013, ApJ, 765, 103
Levinson, A. 2012, ApJ, 756, 174
Levinson, A., & Bromberg, O. 2008, PhRvL, 100, 131101 (LB08)
Mészáros, P., & Rees, M. J. 2000a, ApJ, 541, L5
Mészáros, P., & Rees, M. J. 2000b, ApJ, 530, 292
Mészáros, P., & Waxman, E. 2001, PhRvL, 87, 171102
Murase, K., & Ioka, K. 2013, PhRvL, 111, 121102
Murase, K., Kashiyama, K., & Mészáros, P. 2013, PhRvL, 111, 131102
Ohsuga, K., & Takahashi, H. 2016, ApJ, 818, 162
Pe’er, A., Mészáros, P., & Rees, M. J. 2006, ApJ, 642, 995
Pe’er, A., & Waxman, E. 2004, ApJ, 613, 448
Razzaque, S., Mészáros, P., & Waxman, E. 2003, PhRvD, 68, 083001
Rees, M. J., & Mészáros, P. 1994, ApJ, 430, L93
Ryde, F., Pe’er, A., Nymark, T., et al. 2011, MNRAS, 415, 3693
Sapir, N., Katz, B., & Waxman, E. 2013, ApJ, 774, 79
Sironi, L., & Spitkovsky, A. 2011, ApJ, 726, 75
Stern, B. E. 2003, MNRAS, 345, 590
Stockem, A., Friaza, F., Fonseca, R. A., & Silva, L. O. 2012, ApJ, 755, 68
Svensson, R. 1984, MNRAS, 209, 175
Svensson, R. 1987, MNRAS, 227, 403
Thompson, C., & Gill, R. 2014, ApJ, 791, 46
Thompson, C., & Madau, P. 2000, ApJ, 538, 105
Tolstov, A., Blinnikov, S., Nagataki, S., & Nomoto, K. 2015, ApJ, 811, 47
Tolstov, A. G., Blinnikov, S. I., & Nadyozhin, D. K. 2013, MNRAS, 429, 3181
Vurm, I., & Beloborodov, A. M. 2016, ApJ, 831, 175
Vurm, I., Beloborodov, A. M., & Poutanen, J. 2011, ApJ, 738, 77
Waxman, E., & Loeb, A. 2001, PhRvL, 87, 071101
Weaver, T. A. 1976, ApJS, 32, 233
Yu, H.-F., van Eerten, H. J., Greiner, J., et al. 2015, A&A, 583, A129
Zeldovich, Y. B., & Raizer, Y. P. 1966, Physics of Shock Waves and High-temperature Hydrodynamic Phenomena (New York: Academic)