WOPA Report

Shocks and Particle Shock Acceleration

Marty Lee (Chair), Tony Bell, David Burgess, Ram Cowsik, Tom Intrator, Randy Jokipii, Bob Lin, Christoph Niemann, Anatoly Spitkovsky

A. Importance of Shocks in the Cosmos

Collisions of super-fast-magnetosonic flows occur frequently in the Cosmos from the small-scale flows of the heliosphere to the large-scale flows characteristic of galactic clusters and the jets in Active Galactic Nuclei (AGN). In addition large-amplitude compressive fluctuations evolve to produce shock waves at their compressive edges. Shocks in the solar wind, which creates the heliosphere, include planetary bow shocks, the shocks driven by coronal mass ejections (CMEs), the shocks formed by the collision of fast and slow wind from adjacent regions of the Sun, and the solar wind termination shock. Shocks in the Galaxy and beyond include those driven by supernovae explosions, galactic winds, AGN jets, accretion onto compact objects, galactic motion in galaxy clusters, and Gamma Ray Bursts. These shocks span a huge range of spatial scale, strength and plasma parameter space. Because particle densities are generally very low throughout most of the Cosmos, the mean free path due to Coulomb collisions is typically large compared with the shock spatial scales of interest. Therefore, most shocks are collisionless and the interaction between the far upstream and far downstream plasmas is mediated by electromagnetic fields. An interesting exception is charge-exchange coupling of the interacting plasmas to the atoms present in the partially ionized upstream and downstream plasmas. Despite their diversity, collisionless shocks share common characteristics: they are inferred to accelerate nonthermal particles efficiently and to generate and amplify magnetic fields, in addition to decelerating supersonic flows.

Shocks convert a fraction of the ordered kinetic energy density of the upstream flow to the higher entropy per unit mass downstream flow by dissipative processes occurring in the shock layers. As a result of the nonlinear plasma processes involved in the shock layers, and their diversity depending on the broad range of possible plasma parameters, the physics of shock structure is challenging and remains unclear. The nature of the interaction between the upstream and downstream plasmas depends on (i) the ambient magnetic field and its obliquity relative to the shock normal, (ii) an electric field parallel to the shock normal associated with charge separation in an ion-electron plasma (which can reflect inflowing ions), (iii) a rich variety of streaming instabilities that excite electromagnetic fluctuations, which in turn couple the flows by scattering the individual ions and/or electrons and produce an effective resistivity and viscosity in the shock layers, and (iv) particle acceleration from thermal energies up to relativistic energies. Each of these processes/instabilities, as it affects different particle species including electrons and various ion species, is characterized by its own length scales parallel to the shock normal, so that a particular shock will exhibit multiple length scales depending on which processes/instabilities are relevant.

B. Key Questions

1. Dependence of Shock Structure on Flow Parameters

The mechanisms responsible for the shock structure depend on the parameters of the upstream flow, and these can affect the observable signatures of the shock. However, a systematic understanding of the correspondence between upstream parameters and specific dissipation mechanisms is not available. Since ion and electron transport perpendicular to the magnetic field is suppressed relative to transport parallel to the field, the magnitude and orientation of the upstream magnetic field are critical parameters. At quasi-perpendicular shocks the magnetic field suppresses the interpenetration of the upstream and downstream plasmas, whereas at quasi-parallel shocks streaming instabilities are required to limit the interpenetration. Thus, at quasi-perpendicular shocks a dominant length scale is the proton gyroradius, whereas quasi-parallel shocks have dominant lengthscales of several gyroradii or more. The possible streaming instabilities are many and include the two-stream electrostatic instability, the electromagnetic Weibel instability, the hydromagnetic streaming instability, and possibly the nonresonant firehose instability. The magnetic field magnitude together with the Mach number determine the relative heating of electrons and protons and the critical Mach numbers at which the relative heating changes. The actual heating of the ions occurs at quasi-perpendicular shocks via the reflection of a fraction of the ions at the electrostatic potential (some ions may experience multiple reflections in a process known as shock "surfing"), whereas the heating at quasi-parallel shocks is essentially the first-order Fermi process and the low-energy phase of diffusive shock acceleration (DSA).

2. Are Shocks in the Cosmos Planar and Stationary?

Theoretical models of shocks are often based on the simplifying assumptions that they are stationary, and planar on the length scale of the shock structure. Although the plasma kinetic processes responsible for the dissipation are clearly not planar and stationary, the assumptions are rooted in the idea that the planar stationary "shock" resulting from an average over a large appropriately chosen ensemble of shocks provides a reasonable representation of the shock structure and physics. Community intuition about shock structure is often based on these ensemble-averaged shocks. However, observations and simulations reveal interesting time dependence (sometimes periodic) and important spatial variations along a complex warped surface. Warps with a length scale similar to the turbulence correlation scale in the solar wind have been observed in interplanetary traveling shocks using multi-spacecraft measurements. Warps change the local magnetic obliquity of the shock, which for example affects particle injection and acceleration. A major source of variations along the shock surface is inhomogeneity of the upstream plasma, especially density variations. The density variations create surface warps and inhomogeneous bulk flows downstream, which drive turbulence and magnetic field amplification. An interesting temporally periodic feature of quasi-perpendicular shock structure revealed by simulations is shock re-formation. In general the mechanisms of shock re-formation are unclear. However, for large Mach numbers it appears to result from overstable proton reflection by an unsteady shock potential that results in periodic dissipation and a periodic variation in shock speed and location. Although shock re-formation could have important consequences for particle injection, for example, it is challenging to detect with spacecraft measurements.

3. Particle Injection and Diffusive Shock Acceleration

An important channel of shock dissipation is particle acceleration by a combination of first-order Fermi acceleration and shock drift acceleration known as diffusive shock acceleration (DSA). This mechanism is responsible for most of the energetic particle populations in the heliosphere, the galactic cosmic rays, and presumably many of the nonthermal particle populations in the Cosmos. At higher energies the mechanism is conceptually straightforward, although the nature and excitation of the plasma fluctuations and their impact on particle scattering and transport is not well understood (see Section 4 below). The major uncertainty in application of the mechanism to specific shocks and their associated energetic particles is the rate at which upstream thermal particles are injected into the process. This uncertainty undermines the predictive power of diffusive shock acceleration and is presumably in part responsible for the huge variation in observed ion intensities in solar energetic particle events. The injection rate is certainly dependent on the detailed electromagnetic structure of the shock, which determines the rate at which incoming particles are reflected or scattered back upstream, and it appears to be very sensitive to the local magnetic obliquity. For quasi-perpendicular shocks, particles with energies comparable with the plasma inflow speed are not able to scatter sufficiently to initiate diffusive shock acceleration before being swept through the shock by the magnetic field. Determining the injection mechanism is nontrivial. Even after years of investigations at Earth's bow shock based on ISEE and Cluster data, the origin of the field-aligned beams that initiate the ion acceleration process is unknown. Finally, the lower injection rate of electrons when compared with ions is not well understood, particularly in view of the higher speed of electrons.

4. Amplification of the Magnetic Field at Shocks

The ambient magnetic field fluctuations in the solar wind and interstellar space are generally not sufficient to yield efficient diffusive shock acceleration. However, the accelerating particles are a manifestation of the interpenetrating upstream and downstream plasmas at high energy. The relative streaming of the energetic particles and the upstream flow excites the cyclotron-resonant hydromagnetic streaming instability at lower proton intensities, the non-resonant current-driven instability at higher proton intensities, or variations of these instabilities. The hydromagnetic instability, which maximizes for wave propagation parallel to the ambient magnetic field, is generally evident as an enhancement in the upstream hydromagnetic power at quasi-parallel shocks in interplanetary space, which are able to inject solar wind ions into the acceleration process. The waves often grow to large amplitude and are compressed at the shock, modify the shock structure, and provide effective particle scattering downstream. They also modify the compression ratio sensed by the accelerating particles. Upstream of Earth's bow shock, where wave magnetic amplitudes are comparable with the ambient field strength, the compressive front of a magnetosonic wave is sometimes observed to grow to a Short Large Amplitude Magnetic Structure (SLAMS), which presumably is excited by the free energy released by the ions which it scatters in enhanced numbers back toward the shock. Other compressive wave fronts form "shocklets," which generate whistler precursors. The details of many of these processes are not well understood, particularly the nonlinear evolution of the excited hydromagnetic waves, and need to be pursued by analytical and numerical investigations.

At quasi-perpendicular shocks the streaming instability is not as effective. Particle transport across the average field is primarily by random walk of field lines, which leads to small scattering mean free paths parallel to the shock normal and steep spatial gradients. This configuration is unstable to a version of the Rayleigh-Taylor instability as the upstream plasma is decelerated by the ion pressure gradient. The resulting warped field lines in the shock precursor presumably reduce the magnetic obliquity and increase injection rates and acceleration efficiency. This scenario is speculative and needs to be investigated by further calculations and simulations.

The magnetic field amplification by the non-resonant current-driven instability is now well established observationally, through the analysis of X-ray images of supernova remnants, and theoretically. This result is crucial to the theory for the origin of galactic cosmic rays up to the "knee" at SNR shocks. See Section D.1.

5. What is the Role of Turbulence in Shock Structure?

Much of the material in the four sections above has described the unstable growth of magnetic fluctuations (which are directly coupled to the interpenetrating plasma flows), and their associated velocity fluctuations, density fluctuations and plasma heating, into large amplitude structures as an intrinsic feature of collisionless shocks. This is perhaps not surprising in view of the fact that a shock can generally be viewed as a large-amplitude magnetosonic wave. Although the growth of the magnetic fluctuations may be couched in terms of wave growth and quasilinear theory, the importance of nonlinear wave-particle and wave-wave interactions is apparent in particular for the strong shocks expected in interstellar space. The question arises whether these fluctuations evolve to a turbulent state in which the initial quasilinear associations between velocities and wavevectors are lost to the characteristics of the turbulence? This must certainly be the case for sufficiently strong shocks. This raises the question whether the distinction between quasi-perpendicular and quasi-parallel shocks has any meaning in a turbulent shock? Would a turbulent shock be amenable to a simpler theoretical description?

B. Numerical Simulations of Shocks

Hybrid, Particle-In-Cell, and Test-Particle numerical simulations are certainly powerful tools for studying most of the challenging questions posed in Sections A.1 – 5. There is much current work using these tools to illuminate specific features of shock structure. However, due to limitations in computing power, numerical simulations are always constrained by reduced dimensionality, ignoring electron dynamics, assuming artificially large electron-to-ion mass ratios, reduced spatial domains, reduced duration of the temporal evolution, or other simplifying assumptions. Although unavoidable, these limitations can lead to artificial behavior or difficulties in interpreting the results. For example, in 1-D and 2-D systems particles are tied to their magnetic field line so that their diffusive shock acceleration cannot be accurately described in quasi-perpendicular geometries. Nevertheless, investigating idealized (possibly unrealistic) shock configurations is valuable in order to identify specific processes and instabilities that may play a less clear role in observed shocks.

The most important challenges for numerical simulations are:

- Systematically identify the instabilities relevant to different parameter domains
- Identify injection rates and mechanisms in Mach number/magnetic obliquity space
- Establish the origin(s) of shock re-formation and its effect on particle acceleration
- Investigate the sensitivity of shocks to upstream inhomogeneities
- Simulate the structure of relativistic shocks

C. Laboratory Experiments on Shocks

As direct observations of collisionless shocks in space are limited in our ability of diagnosing and reproducing these phenomena, there is much to learn from controlled laboratory experiments that can provide additional, and reproducible data to benchmark models, and test empirical scaling relations. Numerous experiments on collisionless shocks have been performed in the past four decades. Most of these used Θ-pinches where a shock is generated by imploding B_z, while some used laser-produced plasmas, and a few used combinations of laser-plasmas and ambient gases. The generated shocks had Alfvénic Mach numbers up to ~20 and were generally perpendicular. These experiments often suffered from being too brief, too small in spatial extent, or influenced by the presence of the nearby walls. Pure laser experiments on the other hand could produce high-beta plasmas but were typically not sufficiently magnetized. By the eighties most laboratory experiments on pinch plasmas had completely died out as funding decreased for magnetic-pinch fusion. However, there is still much to learn from laboratory experiments. No laboratory experiment to date has launched perpendicular collisionless shocks over sufficiently large scale-lengths and time-scales for magnetohydrodynamic turbulence to grow and significant ion acceleration to occur. No laboratory experiment has ever been attempted to generate parallel collisionless shocks. In particular particle injection should be studied in laboratory experiments, as it is not only one of the most crucial but at the same time the most poorly understood aspect of DSA theory.

New opportunities to revive laboratory simulations of collisionless shocks arise with the advent of new HED facilities such as the National Ignition Facility (NIF) and better diagnostics capabilities. With adequate magnetization of the target chamber, NIF could deliver enough energy (~500 kJ kinetic) to an exploding plasma to drive high-Mach-number (M_A>10) collisionless shocks over several tens of ion-gyro-periods ($1/\omega_{ci}$) and ion-inertial lengths (c/ω_{pi}). Similarly, a kilojoule-class laser is currently under construction at UCLA's Large Plasma Device, which will be able to drive supercritical (M_A>4) perpendicular shocks in a tenuous (10^{13} cm^{-3}), magnetized, current-free, and quasi-collisionless plasma (λ_{ii} >> size of experiment), that is large enough (20mx0.6m Ø) to support Alfvén waves. These experiments will be able to drive supercritical (M_A>4) perpendicular shocks over several ion-inertial lengths, and could be the first laboratory experiments to generate parallel collisionless shocks. The goal of these experiments is to study shock structure, dissipation, and the effect on the particle velocity distribution in great detail, as well as to study the dynamics of the diamagnetic cavity in the ambient magnetized plasma.

D. Discovery and Exciting Shock Case Studies

1. Supernovae Shocks and the Origin of Galactic Cosmic Rays

The recent X-ray images of SN remnants, direct detection of γ-rays from SN by Cerenkov telescopes like HESS, and theoretical developments on magnetic field amplification at SN shocks essentially validate the zeroth order model of galactic cosmic ray (GCR) acceleration by SNR shocks. However, interesting and challenging questions remain: (i) Are the observed TeV γ-rays produced by ions or electrons, and do we have direct evidence for proton acceleration to the GCR spectrum knee? (ii) What is the GCR composition above the knee, do we need to explain proton acceleration beyond the knee, and if so, how? (iii) How do GCR escape from SN remnants without adiabatic energy loss? (iv) At which stage of SN remnant evolution (probably free-expansion or Sedov) are GCR mostly accelerated, and is there substantial acceleration during the expansion of very young SN remnants into a dense circumstellar medium? (v) Why is the GCR spectrum steeper than can be comfortably explained by momentum-dependent GCR escape from the galaxy? Is this due to nonlinear effects in cosmic-ray-dominated shocks, and, if so, is there evidence for the expected concavity in the spectrum? (vi) Can our understanding of SNR shocks and magnetic field amplification be extended to shocks in AGN and galactic clusters, and can this explain data from AUGER on the directional origin of the highest energy CR? The answers to all of these questions are inter-related and depend primarily on the interaction between GCR and magnetic field in the shock environment. It is clear from simulations that a simple diffusion model does not properly describe the interaction, and detailed observations of particle acceleration at heliospheric shocks show that the picture is indeed much more complicated. Particularly exciting is the possibility of direct comparison of theory with observations of the spatial and temporal variation of non-thermal X-ray emission from SN remnant shocks, Hα observations of precursor conditions, observations of the dependence on shock conditions of the GCR spectrum and energy density, the determination through the spectrum of a hadronic or leptonic origin of γ-rays, and the location by AUGER of the source of the highest energy GCR.

In spite of the apparent success of the SN remnant shock origin of GCRs, a successful theory of GCRs must account for the luminosity of the Galaxy in cosmic rays, their spectra, the spectra of the GCR spallation products, the proximity of their sources, their approximate isotropy and their constancy in time. A theory that accounts for these diverse observations must include a model for the interstellar propagation of GCRs as well. The two questions posed in issue (v) above underscore the coupling between the theory of the GCR source, a model for their interstellar propagation, and the observations. Several questions arise in this context: (i) Can the efficiency of diffusive shock acceleration at a SN remnant shock account for the observed luminosity? (ii) Can the source spectrum inferred from the observed spectrum and the interstellar propagation model be reconciled with the SN remnant shock acceleration theory? [equivalent to issue (v) above] (iii) The secondary to primary GCR ratio as a function of energy shows no evidence of reacceleration of the secondaries. Is that consistent with the observed production of γ-rays in the SN remnants, which implies secondary production near the shock? Why is there no evidence of GCR reacceleration by other SN remnant shocks that happen to be encountered by the diffusing GCR? (iv) Is this theory consistent with the observed constraints on anisotropy and temporal variability? (v) With regard to electrons, why is their spectrum different than that of nuclei? Why does a SN remnant radio spectrum typically turn on only after ~ 300 years? Are there sufficient SN remnants to produce the GCR electrons within ~ 100 pc of Earth as implied by their TeV high-energy rollover? These questions can in part be addressed over the next decade by improved measurements to higher energies of GCR electrons and key secondary to primary ratios.

2. Heliospheric Shocks: Examples in Different Domains of Plasma Parameter Space

Much of the knowledge we have acquired about the variety of possible collisionless shock structures has come from space observations. In the solar wind we have observed shocks with a broad range of Mach numbers and magnetic obliquity. We have observed the complex Earth's bow shock with comparable foreshock and curvature length scales so that phenomena associated with the quasi-perpendicular portion affect the structure of the quasi-parallel portion. In contrast the large-scale interplanetary traveling shocks appear to be locally approximately planar. In the heliocentric radial range 0.1 – 10 AU the variation of the solar wind plasma-β is not large. However, due to the charge-exchange ionization of interstellar hydrogen in the outer heliosphere, the plasma-β of the solar wind increases to ~ 30 at the location of the solar wind termination shock.

The termination shock is dominated by the pressure of the interstellar pickup ions, which do not have a Maxwellian distribution function. It is a unique shock. From the point of view of the “thermal” plasma, it is similar to a hydrodynamic shock. However, it is nearly perpendicular near its “nose”, where the Voyager 1 and 2 spacecraft crossed it so that the ion injection rate is reduced and irregular, and diffusive shock acceleration is strongly ordered by the field orientation. In addition, the magnetic fluctuations upstream and downstream of the shock have a very different character, which is not understood at all. Much remains to be learned about the termination shock, particularly since the pickup proton distribution could not be observed directly by the Voyager spacecraft. IBEX does promise to provide information on the global properties of the shock.

The plasma-β decreases in the inner heliosphere where shocks become more magnetized. However, only the Helios spacecraft in the 1970’s ventured to within 0.4 AU of the Sun. Shocks are expected to play a very important role close to the Sun. They are driven by fast CMEs and are expected to accelerate the highest energy SEPs. They also have very narrow foreshocks due to the small ion gyroradii and are characterized by very short particle acceleration timescales. Their magnetosonic Mach numbers are not large since the Alfven speed can be as large as 500 km/s but they are variable and dependent on the magnetic field morphology near active regions. Shocks may also be generated by rearrangement of the coronal magnetic field and flare energy release, presumably by magnetic reconnection. These shocks may also propagate higher up into the corona. Although they have been detected through type II radio bursts and their line-of-sight density enhancements in coronagraph images, we know very little about coronal shocks including geometry, formation, and onset of particle acceleration. Two exciting opportunities to study shocks near the Sun will be the Solar Orbiter Mission, scheduled to be launched in 2018 and to reach a heliocentric radial distance of 30 solar radii, and Solar Probe, scheduled to be launched in 2018 and to reach 9 solar radii. Although it will be challenging to resolve coronal shock foreshocks, their compression ratios, magnetic obliquity, and particle acceleration characteristics will be measurable.

E. Concluding Statement

In the next decade we should be able to address many of the issues and answer many of the questions posed in this report. The issues/questions and the observational data appear diverse, but there is a strong unity and coherence in the underlying scientific investigation. An alliance between astrophysical observations, heliospheric data, numerical simulation and laboratory experiments can be expected to produce, through mutual support and verification, a convincing picture of shock structure, shock dissipation processes, particle (and as an important case, GCR) acceleration, and magnetic field generation in our heliosphere, our galaxy, and beyond.