Cosmic-Ray Propagation and Interactions in the Galaxy

Andrew W. Strong,1 Igor V. Moskalenko,2 and Vladimir S. Ptuskin3

1Max-Planck-Institut für extraterrestrische Physik, 85741 Garching, Germany; email: aw@mpg.de
2Hansen Experimental Physics Laboratory and Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, California 94305; email: imos@stanford.edu
3Institute for Terrestrial Magnetism, Ionosphere and Radiowave Propagation of the Russian Academy of Sciences (IZMIRAN), Troitsk, Moscow region 142190, Russia; email: vptuskin@izmiran.ru

Key Words
energetic particles, gamma rays, interstellar medium, magnetic fields, plasmas

Abstract
We survey the theory and experimental tests for the propagation of cosmic rays in the Galaxy up to energies of 10^{15} eV. A guide to the previous reviews and essential literature is given, followed by an exposition of basic principles. The basic ideas of cosmic-ray propagation are described, and the physical origin of its processes is explained. The various techniques for computing the observational consequences of the theory are described and contrasted. These include analytical and numerical techniques. We present the comparison of models with data, including direct and indirect—especially γ-ray—observations, and indicate what we can learn about cosmic-ray propagation. Some important topics, including electron and antiparticle propagation, are chosen for discussion.
1. INTRODUCTION

Cosmic rays (CRs) are almost unique in astrophysics in that they can be sampled directly, not just observed via electromagnetic radiation. Other examples are meteorites and stardust. CRs provide us with a detailed elemental and isotopic sample of the current (few million years old) interstellar medium (ISM) that is not available in any other way. This makes the subject especially rich and complementary to other disciplines.

CRs have been featured in approximately 15 articles from 1952 to 1989 in the Annual Reviews series. These reviews address heavy nuclei (1), collective transport effects (2), composition (3), and propagation (4). Cox’s recent review, The Three-Phase Interstellar Medium (5), contains much discussion of CRs as one essential component of the ISM, but inevitably no mention of their propagation. Two recent Annual Review articles (6, 7) extensively discuss the relation of CRs to turbulence; thus, we do not try to cover this.
A basic reference is the book *Astrophysics of Cosmic Rays* (8), which expounds all the essential concepts and is an update of the classic, *The Origin of Cosmic Rays* (9), which laid the modern foundations of the subject, with an updated presentation in Reference 10. Good books for basic expositions include References 11 and 12, and for high energies, Reference 13. A basic text emphasizing theory is Reference 14, while the book *Astrophysics of Galactic Cosmic Rays* (15) gives a valuable overview of the experimental data and theoretical ideas as of 2001. The biannual International Cosmic Ray Conference proceedings (many of which can be found at [http://adswww.harvard.edu/](http://adswww.harvard.edu/)) are also an essential source of information, especially for the latest news on the subject.

Recently, a plethora of reviews have appeared on the subject of CRs above $10^{15}$ eV (e.g., 16–18); on interactions (19); on experiments and astrophysics (20); and more on astrophysics, propagation, and composition (21–23); in addition to a review of models (24). We recommend Reference 25 and the very up to date References 26 and 27. Therefore, this topic has been excluded here. At the lowest-energy end, we note that MeV particles are nonthermal (even if not relativistic) and must be mentioned in a review of CRs, especially because they are important sources of heating and ionization of the ISM (28). As one example of their far-ranging influence, star formation in molecular clouds may be suppressed by CRs produced in supernova remnants (SNRs) nearby (29).

It is worth distinguishing between two ways of approaching CR propagation: either from the particle point of view, including the spectrum and interactions, or treating the CRs as a weightless, collisionless relativistic gas with pressure and energy and considering it alongside other components of the ISM (30, 31). Both ways of looking at the problem are valid up to a point, but for consistency, a unified approach is desirable and, to our knowledge, has never been attempted. The nearest approaches to this are in References 32 and 33. Most papers exclusively address one or the other aspect. The first approach is required for comparison with observations of CRs (direct and indirect), whereas the second is required for the ISM: stability, heating, and so on (5).

The major recent advances in the field are the high-quality measurements of isotopic composition and element spectra, and observations by γ-ray telescopes, both satellite and ground based. Space does not allow discussion of the observational data here, but the figures give an illustrative overview of what is now available from both direct (Figures 1–13) and indirect (γ-ray) measurements (Figures 14–16). Concerning the origin of CRs, we follow Cesarsky’s review: “We will, for the most part, sidestep this problem” (4). Hence, we omit CR sources, including composition and acceleration; for SNRs as CR sources, the literature can be traced back from the most recent High Energy Stereoscopic System (H.E.S.S.) TeV γ-ray results (34). We also omit solar modulation, Galaxy clusters, and extragalactic CRs. We restrict our attention mostly to our own Galaxy, but mention important information coming from external galaxies (via synchrotron radiation).

We first introduce the theoretical background, and then consider the confrontation of theory with observation. A number of particular topics are selected for further discussion.
Figure 1

Compilation of spectral data $10^{10} - 10^{17}$ eV for proton, helium, and iron, combining balloon, satellite, and ground-based measurements. Figure adapted with permission from Reference 18 and G. Hörandel (private communication).
2. COSMIC-RAY PROPAGATION: THEORY

2.1. Basics and Approaches

We present the basic concepts of CR propagation and techniques for relating these to observational data. Practically all our knowledge of CR propagation comes via secondary CRs, with additional information from $\gamma$-rays and synchrotron radiation. Note why secondary nuclei in particular are a good probe of CR propagation: The fact that the primary nuclei are measured (at least locally) means that the secondary production functions can be computed from primary spectra, cross sections, and interstellar gas densities with reasonable precision. The secondaries can then be propagated and compared with observations.

Since the realization that CRs fill the Galaxy, it has been clear that nuclear interactions imply that their composition contains information on their propagation (44). A historical event was the arrival of satellite measurements of isotopic Li, Be, and B in the 1970s (45). Since then the subject has expanded enormously, with models of increasing degrees of sophistication. The simple observation that the observed composition of CRs is different from that of solar elements in that rare solar-system nuclei such as B are abundant in CRs proves the importance of propagation in the

Secondary cosmic rays: cosmic rays generated by spallation of primary cosmic rays on interstellar gas
Figure 3

(a) Compilation of spectral data for element groups CNO, NeMgSi, and Fe (36) from HEAO-3, SANRIKU, CRN, SOKOL, JACEE, and RUNJOB. Figure adapted with permission from the American Astronomical Society. (b) Compilation of separate even-Z elements from preliminary ATIC-2, HEAO-3, CRN, and TRACER. The ATIC-2 data suggest a hardening above 10 TeV. Plot taken from the ATIC Collaboration (35). Figure adapted with permission from Reference 35.
2.2. Propagation Equation

The CR propagation equation for a particular particle species can be written in the general form

$$\frac{\partial \psi(\vec{r}, p, t)}{\partial t} = q(\vec{r}, p, t) + \nabla \cdot (D_{xx} \nabla \psi - \vec{V} \psi)$$

$$+ \frac{\partial}{\partial p} p^2 D_{pp} \frac{1}{p^2} \psi - \frac{\partial}{\partial p} \left[ \frac{\dot{p} \psi - p}{3} (\vec{V} \cdot \nabla) \psi \right] - \frac{1}{\tau_f} \psi - \frac{1}{\tau_r} \psi,$$

where $\psi(\vec{r}, p, t)$ is the CR density per unit of total particle momentum $p$ at position $\vec{r}$, $\psi(p)dp = 4\pi p^2 f(\vec{p})dp$ in terms of phase-space density $f(\vec{p})$; $q(\vec{r}, p, t)$ is the source term including primary, spallation, and decay contributions; $D_{xx}$ is the spatial...
Figure 5
Measurements of the positron spectrum, including data from MASS91, AMS01, CAPRICE94, and HEAT. Propagation calculations for interstellar (dotted curves) and modulated (solid curves) spectra are shown. Red lines, GALPROP (39); orange lines, Reference 40. Figure adapted from Reference 38 with permission from Astron. Astrophys.

GALPROP: Galactic Propagation (code)

β: velocity divided by speed of light

Diffusion coefficient; \( \vec{V} \) is the convection velocity; diffusive reacceleration is described as diffusion in momentum space and is determined by the coefficient \( D_{pp} \); \( \dot{p} = dp/dt \) is the momentum gain or loss rate; \( \tau_f \) is the timescale for loss by fragmentation; and \( \tau_r \) is the timescale for radioactive decay.

CR sources are usually assumed to be concentrated near the galactic disk and to have a radial distribution similar to, for example, SNRs. A source injection spectrum and its isotopic composition are required. Usually composition is initially based on solar photospheric elements, but can be determined iteratively from the CR data for later comparison with solar composition. The spallation part of \( q(\vec{r}, \beta, p/Z) \) depends on all progenitor species and their energy-dependent cross sections, and the gas density \( n(\vec{r}) \). In general, it is assumed that the spallation products have the same kinetic energy per nucleon as the progenitor. K-electron capture and electron stripping can be included via \( \tau_f \) and \( q \). \( D_{ee} \) is, in general, a function of \( (\vec{r}, \beta, p/Z) \), where \( \beta = v/c \), \( Z \) is the charge, and \( p/Z \) determines the gyroradius in a given magnetic field. \( D_{ee} \) may be isotropic, or more realistically anisotropic, and may be influenced by the CRs (e.g., in wave-damping models). \( D_{pp} \) is related to \( D_{ee} \) by \( D_{pp}/D_{ee} \propto p^2 \), with the proportionality constant dependent on the theory of stochastic reacceleration (8, 46), as described in Section 2.5. \( \vec{V} \) is a function of \( \vec{r} \) and depends on the nature of the galactic wind. The term in \( \nabla \cdot \vec{V} \) represents adiabatic momentum gain or loss in the nonuniform flow of gas, with a frozen-in magnetic field whose inhomogeneities scatter the CRs. \( \tau_f \) depends on the total spallation cross section and \( n(\vec{r}) \). \( n(\vec{r}) \) can be based on surveys of atomic and molecular gas, but can also incorporate small-scale
Figure 6

(a) Measurements of the electron and positron spectrum by AMS01 and HEAT. Figure adapted with permission from Reference 41. (b) \( \frac{\psi^+}{(\psi^+ + \psi^-)} \) ratio, including HEAT and AMS01 (42). For the ratio, a propagation calculation (solid line) (39), possible contributions from pulsars (dashed line), and dark matter annihilations (dotted line) are shown. Figure adapted with permission from Phys. Rev. Lett. 93:241102 by Beatty et al. Copyright (2004) by the American Physical Society.

variations such as the region of low gas density surrounding the Sun. The presence of interstellar helium at approximately 10% of hydrogen by number must be included. Heavier components of the ISM are not important for producing CRs by spallation. This equation treats continuous momentum loss only. Catastrophic losses can be included via \( \tau_f \) and \( q \). CR electron, positron, and antiproton propagation constitute special cases of this equation, differing only in their energy losses and production rates. The boundary conditions depend on the model. Often, \( \psi = 0 \) is assumed at the halo boundary, where particles escape into intergalactic space, but this is obviously just an approximation (because the intergalactic flux is not zero) that can be relaxed for models with a physical treatment of the boundary.

Equation 1 is a time-dependent equation. Usually the steady-state solution is required, which can be obtained by either setting \( \frac{\partial \psi}{\partial t} = 0 \) or following the time dependence until a steady state is reached. The latter procedure is much easier to implement numerically. The time dependence of \( q \) is neglected unless effects of
nearby recent sources or the stochastic nature of sources is being studied. By starting with the solution for the heaviest primaries and using this to compute the spallation source for their products, the complete system can be solved, including secondaries, terciaries, and so on. Then the CR spectra at the solar position can be compared with direct observations, including solar modulation if required.

Source abundances are determined iteratively, comparing propagation calculations with data. For nuclei with very small source abundances, the source values are masked by secondaries and cross-section uncertainties and are therefore difficult to determine. Webber (47) gives a ranking of easy to impossible for the possibility of getting the source abundances using the Advanced Composition Explorer (ACE) data. A recent review of the high-precision abundances from the ACE is in Reference 48, and for Ulysses, in Reference 49. For a useful summary of the various astrophysical abundances relevant to interpreting CR abundances, see Reference 50.

2.3. Diffusion

The concept of CR diffusion explains why energetic charged particles have highly isotropic distributions and why they are well retained in the Galaxy. The galactic magnetic field that tangles the trajectories of particles plays a crucial role in this process. Typical values of the diffusion coefficient found from fitting to CR data are \( D_x \sim (3 - 5) \times 10^{28} \text{ cm}^2 \text{ s}^{-1} \) at energy \( \sim 1 \text{ GeV/n} \), and increase with magnetic rigidity as \( R^{0.3} - R^{0.6} \) in different versions of the empirical diffusion model of CR propagation.

On the microscopic level, the diffusion of CRs results from particle scattering on random magnetohydrodynamic (MHD) waves and discontinuities. MHD waves arise
in magnetized plasmas in response to perturbations. The effective collision integral for charged energetic particles moving in a magnetic field with small random fluctuations \( \delta B \ll B \) can be taken from the standard quasi-linear theory of plasma turbulence (51). The wave-particle interaction is of resonant character, so that an energetic particle is scattered predominantly by those irregularities of magnetic field that have their projection of the wave vector on the average magnetic field direction equal to 

\[
k_{11} = \frac{\pm s}{r_g \mu},
\]

where \( r_g = \beta c / Ze B \) is the particle gyroradius and \( \mu \) is the pitch angle. The integers \( s = 0, 1, 2 \ldots \) correspond to cyclotron resonances of different orders. The efficiency of scattering depends on the polarization of the waves and on their distribution in the \( k \)-space of the wave vectors, where \( k \) is the wave number. The first-order resonance \( s = 1 \) is the most important for the isotropic and also for the one-dimensional distribution of random MHD waves along the average magnetic field. In some cases—for calculation of scattering at small \( \mu \) and of perpendicular diffusion—the broadening of resonances and magnetic mirroring effects should be taken into account. The resulting spatial diffusion is strongly anisotropic locally and goes predominantly along the magnetic field lines. However, strong fluctuations of the magnetic field on large scales \( L \sim 100 \, \text{pc} \), where the strength of the random field is several times higher than the average field strength, lead to the isotropization of global CR diffusion in the Galaxy. The rigorous treatment of this effect is not trivial, as the field is almost static and the strictly one-dimensional diffusion along the magnetic field lines does not lead to nonzero diffusion perpendicular to the field (see Reference 52 and references cited therein).

Following several detailed reviews of the theory of CR diffusion (8, 14, 53, 54), the diffusion coefficient at \( r_g < L \) can be roughly estimated as

\[
D_{xx} \approx (\delta B_{\text{res}} / B)^{-1} \frac{v_{\text{rg}}}{3},
\]

where \( \delta B_{\text{res}} \) is the amplitude of the random field at the resonant wave number \( k_{\text{res}} = 1 / r_g \). The spectral energy density of interstellar turbulence has a power-law form

\[ w(k)dk \sim k^{-2+\delta}dk, \]

for a wide range of wave numbers \( 1/(10^{20} \, \text{cm}) \leq k < 1/(10^{3} \, \text{cm}) \) (see Reference 6), and the strength of the random field at the main scale is \( \delta B \approx 5 \, \mu \text{G} \). This gives an estimate of the diffusion coefficient

\[
D_{xx} \approx 2 \times 10^{21} \beta B_{\text{GV}}^{1/3} \text{cm}^2 \text{s}^{-1}
\]

for all CR particles with magnetic rigidities \( R < 10^8 \, \text{GV} \), in fair agreement with the empirical diffusion model (the version with distributed reacceleration). The scaling law \( D_{xx} \sim R^{1/3} \) is determined by the value of the exponent \( a = 1/3 \), typical for a Kolmogorov spectrum. Theoretically, the Kolmogorov-type spectrum may refer only to some part of the MHD turbulence that includes the (Alfvénic) structures strongly elongated along the magnetic field direction and that cannot provide the significant scattering and required diffusion of CRs (55). In parallel, the more isotropic (fast magnetosonic) part of the turbulence, with a smaller value of the random field at the main scale and with the exponent \( a = 1/2 \) typical for the Kraichnan-type turbulence spectrum, may exist in the ISM (56). The Kraichnan spectrum gives a scaling

\[
D_{xx} \sim R^{1/2},
\]

which is close to the high-energy asymptotic form of the diffusion coefficient obtained in the plain diffusion version of the empirical propagation model. Thus, the approach based on kinetic theory gives a proper estimate of the diffusion coefficient and predicts a power-law dependence of diffusion on magnetic rigidity. However, the actual diffusion coefficient must be determined with the help of empirical models of CR propagation in the Galaxy.
2.4. Convection

Although the most frequently considered mode of CR transport is diffusion, the existence of galactic winds in many galaxies suggests that convective (or advective) transport may be important. Winds are common in galaxies and can be CR driven (57). CRs play a dynamical role in galactic halos (58, 59). Convection not only transports CRs, but can also produce adiabatic energy losses as the wind speed increases away from the disk. Convection was first considered in Reference 60 and followed up in References 61–65. Both one-zone and two-zone models have been studied: a one-zone model has convection and diffusion everywhere; a two-zone model has diffusion alone up to some distance from the plane, and diffusion plus convection beyond.

Surprisingly, a recent review on galactic winds (66) hardly mentions CRs. In the same volume, Cox's review (5) includes CRs as a basic component. Direct evidence for winds in our own Galaxy seems to be confined to the galactic center region from X-ray images. However, Cox is not sure there is a wind: “A Galactic wind may be occurring, but I do not believe that it carries off a significant fraction of the supernova power from the Solar Neighborhood because it would carry off a similar power in the pervading cosmic rays...” (p. 370).

For one-zone diffusion-convection models, a good diagnostic is the energy dependence of the secondary-to-primary ratio. A purely convective transport would have no energy dependence (apart from the velocity dependence of the reaction rate), contrary to what is observed. If the diffusion rate decreases with decreasing energy, any convection will eventually take over and cause the secondary-to-primary ratio to flatten at low energy. This is observed, but convection (proposed in Reference 64 to explain just this effect) does not reproduce, for example, the CR boron-to-carbon ratio (B/C) very well (67). Another test is provided by radioactive isotopes, which effectively constrain the wind speed to <10 km s\(^{-1}\) kpc\(^{-1}\) for a speed that increases linearly with distance from the disk (67). A value of \(\approx 15\) km s\(^{-1}\) (constant speed wind) is required to fit B/C even in the presence of reacceleration, according to Reference 68, which can be compared to 30 km s\(^{-1}\) in the wind model of Reference 69. The latter value implies an energy dependence of the diffusion coefficient that may conflict with CR anisotropy.

Ptuskin et al. (33) studied a self-consistent two-zone model with a wind driven by CRs and thermal gas in a rotating Galaxy. The CR propagation is entirely diffusive in a zone |z| < 1 kpc, and diffusive-convective outside. CRs reaching the convective zone do not return, so they act as a halo boundary—with height varying with energy and Galactocentric radius (the distance of a point from the galactic center). It is possible to explain the energy dependence of the secondary-to-primary ratio with this model, and it is also claimed that it is consistent with radioactive isotopes. The effect of a galactic wind on the radial CR gradient has been investigated in Reference 70, in which the authors constructed a self-consistent model with the wind driven by CRs and with anisotropic diffusion. The convective velocities involved in the outer zone are large (100 km s\(^{-1}\)), but this model is still consistent with radioactive CR nuclei, which set a much lower limit (67) because this limit is applicable only in the
inner zone. Observational support of such models would require direct evidence for a galactic wind in the halo.

2.5. Reacceleration

In addition to spatial diffusion, the scattering of CR particles on randomly moving MHD waves leads to stochastic acceleration, which is described in the transport equation as diffusion in momentum space with some diffusion coefficient $D_{pp}$. One can estimate it as $D_{pp} = p^2 V_a^2 / (9 D_{xx})$, where the Alfvén velocity $V_a$ is introduced as a characteristic velocity of weak disturbances propagating in a magnetic field (see References 8 and 14 for rigorous formulas).

Distributed acceleration in the entire galactic volume cannot serve as the main mechanism of acceleration of CRs, at least in the energy range $1 - 100$ GeV/n. In this case, the particles of higher energy would spend a longer time in the system, which would result in an increase of the relative abundance of secondary nuclei as energy increases, contrary to observation. This argument does not hold at low energies, where distributed acceleration may be strong, and it may explain the existence of peaks in the ratios of secondary-to-primary nuclei at approximately 1 GeV/n if the distributed acceleration becomes significant at this energy. The process of distributed acceleration in the ISM is also referred to as reacceleration, to distinguish it from the primary acceleration process that occurs in the CR sources. It has been shown (46, 71) that the observed dependence of abundance of secondary nuclei on energy can be explained in the model with reacceleration (a) if the CR diffusion coefficient varies as a single power law of rigidity $D_{xx} \sim R^a$, with an exponent $a \sim 0.3$ over the whole energy range (corresponding to particle scattering on MHD turbulence with a Kolmogorov spectrum) and (b) if the Alfvén velocity is $V_a \sim 30$ km s$^{-1}$, which is close to its actual value in the ISM.

In addition to stable secondary nuclei, the secondary K-capture isotopes are useful for the study of possible reacceleration in the ISM (72). The isotopes $^{37}$Ar, $^{44}$Ti, $^{49}$V, and $^{51}$Cr and some others decay rapidly by electron capture at low energies, where energetic ions can have an orbital electron. The probability of having an orbital electron depends strongly on energy, and because of this the abundance of these isotopes and of their decay products are strong functions of energy and sensitive to changes of particle energy in the ISM. The first measurements of an energy-dependent decay of $^{49}$V and $^{51}$Cr in CRs (73) were used to test the rate of distributed interstellar reacceleration (74), but refinement of nuclear production cross sections is required to draw definite conclusions.

The gain of particle energy in the process of reacceleration is accompanied by a corresponding energy loss of the interstellar MHD turbulence. According to calculations (75), the dissipation on CRs may significantly influence the Kraichnan nonlinear cascade of waves at less than $10^{13}$ cm, and may even terminate the cascade at small scales. This results in a self-consistent change of the rigidity dependence of the diffusion coefficient, with a steep rise of $D_{xx}$ to small rigidities. The scheme explains the high-energy scaling of diffusion $D_{xx} \sim R^{0.5}$ and offers an explanation for the observed energy dependence of secondary-to-primary ratios.
As mentioned above, the data on secondary nuclei provide evidence against strong reacceleration in the entire Galaxy at 1–100 GeV/n. However, the spectra of secondaries can be modified considerably owing to processes in the source regions, with a small total galactic volume filling factor, for the regions where the high-velocity SNR shocks accelerate primary CRs. Two effects could be operating there, and both lead to the production of a component of secondaries with flat energy spectra (76, 77). One effect is the production of secondaries in SNRs by the spallation of primary nuclei that have a flat source energy spectrum close to $E^{-2}$. Another effect is the direct acceleration by strong SNR shocks of background secondary nuclei residing in the ISM; again the secondaries acquire the flat source energy spectrum. Calculations (77) showed that these effects may produce a flat component of secondary nuclei rising above the standard steep spectrum of secondaries at energies above approximately 100 GeV/n.

### 2.6. Galactic Structure

Almost all aspects of galactic structure affect CR propagation, but the most important are the gas content for secondary production and the interstellar radiation field (ISRF) and magnetic field for electron energy losses. The magnetic field is clearly also important for diffusion, but the precise absolute magnitude and large-scale structure are less important (at least for CRs below $10^{15}$ eV) than the turbulence properties.

The distribution of atomic hydrogen is reasonably well known from 21-cm surveys, but the distribution of molecular hydrogen is less well known because it can only be estimated using the CO molecular tracer; the associated scaling factor is hard to determine and may depend on its position in the Galaxy. In fact CR-gas interactions provide one of the best methods for determining the molecular hydrogen content of the Galaxy because of its basic simplicity, as we describe in the section on $\gamma$-rays. For more details, refer to Reference 78.

The galactic magnetic field can be determined from pulsar rotation and dispersion measures combined with a model for the distribution of ionized gas. A large-scale field of a few microgauss aligned with spiral arms exists, but there is no general agreement on the details (79). One recent analysis gives a bisymmetric model for the large-scale galactic magnetic field with reversals on arm-interarm boundaries (80). Independent estimates of the strength and distribution of the field can be made by simultaneous analysis of radio synchrotron, CRs, and $\gamma$-ray data, and these confirm a value of a few microgauss, increasing toward the inner Galaxy (81). Much effort has gone into constructing magnetic field models to study propagation at energies $>10^{15}$ eV, where the Larmor radius is large enough for the global topology to be important. This is relevant to CR anisotropy and the search for point sources. Because this is excluded from our review, we refer the reader to References 82–84.

The ISRF comes from stars of all types and is processed by absorption and re-emission by interstellar dust. It extends from the far-infrared, through the optical, to the ultraviolet. Computing the ISRF is difficult, but a great deal of new information on the stellar content of the Galaxy and dust is now available to make better models.
for use in propagation codes (85). The other important radiation field is the cosmic microwave background, with its well-known black-body spectrum.

The local environment around the Sun (86) is also important. For example, the Local Bubble can affect the radioactive nuclei, as described in Section 3.2. Extensive coverage of the local environment including CRs is in Reference 87.

### 2.7. Interactions

This large subject is well covered in the literature. Details of the essential processes with references have been conveniently collected in our series of papers: energy losses of nuclei and electrons (67); bremsstrahlung and synchrotron emission (81); inverse-Compton emission including anisotropic scattering (88); and pion production of $\gamma$-rays, electrons, and positrons (39). Pion production was recently studied in great detail using modern particle-physics codes (89, 90), giving spectra harder by 0.05 in the index and $\gamma$-ray yields somewhat higher at a few GeV than older treatments. New, more accurate parameterizations will be important for the new generation of CRs and $\gamma$-ray experiments. Reference 91 and references therein are useful guides to spallation cross-section measurements and models. A summary of recent advances appear in References 92–94. Accounts of radioactive and K-capture processes can be found in References 73 and 95–98.

### 2.8. Weighted Slabs and Leaky Boxes

As mentioned at the start of this section, at present we believe that the diffusion model with possible inclusion of convection provides the most adequate description of CR transport in the Galaxy at energies below approximately $10^{17}$ eV. The closely related leaky-box and weighted-slab formalisms have provided the basis for most literature interpreting CR data.

The leaky-box model uses the simple picture of particles injected by sources distributed uniformly over some volume (box) and escaping from this volume with an escape time independent of position, and a uniform distribution of gas and radiation fields. The leaky-box equations express this mathematically with source and leakage terms. In the leaky-box model, the diffusion and convection terms are approximated by the leakage term, with some characteristic escape time of CRs from the Galaxy. The escape time $\tau_{\text{esc}}$ may be a function of particle energy (momentum), charge, and mass number if needed, but it does not depend on the spatial coordinates. There are two models where the leaky-box equations can be obtained as a correct approximation to the diffusion model: (a) the model with fast CR diffusion in the Galaxy and particle reflection at the CR halo boundaries with some probability of escape (9), and (b) the flat-halo model ($z_h \ll R$) with thin source and gas disks ($z_{\text{gas}} \ll z_h$), which is equivalent formally to the leaky-box model in the case where stable nuclei are considered (10). The nuclear fragmentation is actually determined not by the escape time $\tau_{\text{esc}}$, but rather by the escape length in g cm$^{-2}$: $x = \nu \rho \tau_{\text{esc}}$, where $\rho$ is the average gas density of interstellar gas in a Galaxy with CR distributed over the whole galactic volume, including the halo.

$z_h$: height of cosmic ray halo in direction perpendicular to galactic plane.
The solution of a system of coupled transport equations for all isotopes involved in the process of nuclear fragmentation is required for studying CR propagation. A powerful method, the weighted-slab technique, which consists of splitting the problem into astrophysical and nuclear parts, was suggested for this problem (9, 99) before the modern computer epoch. The nuclear-fragmentation problem is solved in terms of the slab model, wherein the CR beam is allowed to traverse a thickness \( x \) of the interstellar gas and these solutions are integrated over all values of \( x \), weighted with a distribution function \( G(x) \) derived from an astrophysical propagation model. In its standard realization (100, 101), the weighted-slab method breaks down for low-energy CRs, where one has strong energy dependence of nuclear cross sections, strong energy losses, and energy-dependent diffusion. Furthermore, if the diffusion coefficient depends on the nuclear species, the method has rather significant errors. After some modification (102), the weighted-slab method becomes rigorous for the important special case of separable dependence of the diffusion coefficient on particle energy (or rigidity) and position, with no convective transport. The modified weighted-slab method was applied to a few simple diffusion models in References 69 and 74. The weighted-slab method can also be applied to the solution of the leaky-box equations. It can easily be shown that the leaky-box model has an exponential distribution of path lengths \( x \) (grammage) \( G(x) \propto \exp(-x/X) \), with the mean grammage equal to the escape length \( X \).

In a purely empirical approach, one can try to determine the shape of the distribution function \( G(x) \) that best fits the data on abundances of stable primary and secondary nuclei (1). The shape of \( G(x) \) is close to exponential: \( G(x) \propto \exp(-x/X(R, \beta)) \), where \( \beta = \text{velocity of particle/velocity of light} \), and this justifies the use of the leaky-box model in this case. There are several recent calculations of \( G(x) \) (69, 74, 103, 104).

The possible existence of truncation, a deficit at small path lengths (below a few grams per centimeters squared at energies near 1 GeV/n), relative to an exponential path-length distribution, has been discussed for decades (1, 101, 105, 106). The problem was not solved mainly owing to cross-section uncertainties. In a consistent theory of CR diffusion and nuclear fragmentation in the cloudy ISM, the truncation occurs naturally if some fraction of CR sources resides inside dense giant molecular clouds (107).

For radioactive nuclei, the classical approach is to compute the surviving fraction, which is the ratio of the observed abundance to that expected in the case of no decay. Often the result is given in the form of an effective mean gas density, to be compared with the average density in the Galaxy, but this density should not be taken at face value. The surviving fraction can be better related to physical parameters (108). None of these methods can face the complexities of propagation of CR electrons and positrons with their large energy and spatially dependent energy losses.

### 2.9. Explicit Models

Finally, the mathematical effort required to put the three-dimensional Galaxy into a one-dimensional formalism becomes overwhelming, and it seems better to work...
in physical space from the beginning. This approach is intuitively simple and easy to interpret. We can call these explicit solutions. The explicit solution approach including secondaries was pioneered in Reference 10, and its application to newer data is described in References 65 and 109, with analytical solutions for two-dimensional diffusion-convection models with a CR source distribution—which, however, had many restrictive approximations to make them tractable (no energy losses, simple gas model). More recently, a two-dimensional semi-empirical model that includes energy losses and reacceleration was developed (68, 110). This is a closed-form solution expressed as a Green’s function to be integrated over the sources. It incorporates a radial CR source distribution, but the gas model is a simple constant density within the disk. Reference 111 gives an analytical solution for the time-dependent case with a generalized gas distribution but without energy losses. (This shows again the problem of handling both gas and energy losses simultaneously in analytical schemes.)

A myriad sources model (112), which is actually a Green’s function method without energy losses, yields similar results to Reference 113 for the diffusion coefficient and halo size. However, applying the no-energy loss case to ACE data is not justified, and authors have pointed out some defects in their formulation (111). A three-dimensional analytical propagation method has been developed (114, 115) with energy loss and reacceleration, going via a path-length distribution, but it cannot handle ionization losses correctly (see section 2.2 in Reference 115). They have no spatial boundaries and rather simplified (exponential) forms for the gas and other distributions. An approach adapted to fine-scale spatial and temporal variations has been described (116). This uses a Green’s function without energy losses or a detailed gas model and hence is limited in its application, but is useful for studying the effect of discrete sources.

The most advanced, explicit solution to date is the fully numerical model described in the next section. Even this has limitations in treating some aspects (e.g., when particle trajectories become important at high energies), so one may ask whether a fully Monte Carlo approach (as is commonly used for energies \( > 10^{15} \) eV) would be better in the future, given increasing computing power. This would allow effects such as field-line diffusion (important for propagation perpendicular to the galactic plane) to be explicitly included. However, it is still challenging. A particle at GeV energies diffusing with a mean free path of 1 pc in a Galaxy with kpc-size halo height takes \( \approx 10^6 \) scatterings to leave the Galaxy, which even now would need supercomputers to obtain adequate statistics. Hence, we expect the numerical solution of the propagation equations to remain an important approach for the foreseeable future.

2.10. GALPROP

The Galactic Propagation (GALPROP) code (67) was created with the following aims: (a) to enable simultaneous predictions of all relevant observations, including CR nuclei, electrons, and positrons; \( \gamma \)-rays; and synchrotron radiation; (b) to overcome the limitations of analytical and semi-analytical methods, taking advantage of advances in computing power as CR, \( \gamma \)-ray, and other data become more accurate; (c) to incorporate current information on galactic structure and source distributions; and (d) to provide a publicly available code as a basis for further expansion. The first
The computing resources required by GALPROP are moderate by today's standards. Although GALPROP has the ambitious goal of being realistic, it is obvious that any such model can be only a crude approximation to reality. Some known limitations are that it applies only to energies below $10^{15}$ eV (there are no trajectory calculations),
it is limited to scales >10 pc (no clumpy ISM; limited by computer power), and the magnetic field is treated as random for synchrotron emission whereas the regular component affects the structure of radio emission. For the cases where these limitations apply, other techniques may be more appropriate and they provide a goal for future developments of GALPROP.

2.11. Numerical versus Analytical

The following expresses the authors’ opinion on this matter. The analytical approaches claim to have various advantages as follows:

1. Physical insight: Of course analytical solutions for simple cases give insight into the relations between the quantities involved, and are useful for rough estimates. In fact, the analytical formulae may become so complicated that no insight is gained. In contrast, the numerical models are very intuitive because they explicitly generate the CR distribution over the Galaxy for all species.

2. Equivalence to full solution of propagation equation: This is true only under restrictive conditions, especially involving energy losses and spatially varying densities. Electrons and positrons are beyond analytical methods because their energy losses are spatially dependent and different processes are important in different energy ranges, whereas these particles are an essential component of the CRs.

3. Faster, easier to compute: With today’s computers, the issue of speed has become irrelevant, and the implementation of a numerical model is no more difficult than the complicated integrals over Bessel functions and so on.

In summary, we can do no better than a quotation from a paper of 26 years ago:

It is unclear whether one would wish to go much beyond the generalizations discussed above for an analytically soluble diffusion model. The added insight from any analytic solution over a purely numerical approach is quickly cancelled by the growing complexity of the formulae. With rapidly developing computational capabilities, one could profitably employ numerical solutions . . .

J.M. Wallace: *Galactic Cosmic-Ray Diffusion with Arbitrary Radial Distributions* (118)

For CR air-shower calculations, analytical methods gave way to numerical ones at least 40 years ago.

2.12. Self-Consistent Models

A few attempts at a self-consistent description of CRs in the Galaxy have been made by including them as a relativistic gas as one component of ISM dynamics. This is obviously much more difficult than the phenomenological models described above, which treat propagation in a prescribed environment. In References 32 and 33, three-dimensional models of the magnetized ISM with a CR-driven wind were made. It is claimed that these models are also consistent with CR secondary-to-primary ratios.
Such a wind has been put forward as a possible explanation of the CR-gradient problem (70). The Parker instability has been reanalyzed recently (119) using anisotropic diffusion (120) and was followed by a CR-driven galactic dynamo model (121), which used an extension of the ZEUS—three-dimensional MHD code (30) including CR propagation and sources. CR propagation in a magnetic field produced by dynamo action of a turbulent flow (31) presents the whole subject from a novel viewpoint. The extension of such approaches to include CR spectra, secondaries, γ-rays, and so on—which would provide a complete set of comparisons with observations—would be very desirable but has not yet been attempted. Another kind of self-consistency is to include the effect of CRs on the diffusion coefficient (75), as described in more detail in Section 2.5.

3. CONFRONTATION OF THEORY WITH DATA

3.1. Stable Secondary-to-Primary Ratios

The reference ratio is almost always B/C because B is entirely secondary. The measurements are better than for other ratios and are available up to 100 GeV. Because C, N, and O are the major progenitors of B, the production cross sections are better known than, for example, in the cases of Be and Li (94, 122).

The usual procedure is to use a leaky-box or weighted-slab formalism with the empirical rigidity dependences $X(R) = (\beta/\beta_0)X_0$ and $(\beta/\beta_0)(R/R_0)^{-\alpha}X_0$ for $R < R_0$ and $R > R_0$, respectively. The break at $R_0$ is required because B/C is observed to decrease to low energies faster than the $\beta$ dependence (which just describes the velocity effect on the reaction rate). The source composition depends on the form and parameters of $X(R)$, and vice versa, (because, for example, B is produced by C, N, and O, etc.), so the procedure is iterative, starting from a solar-like composition.

A typical parameter set fitting the data (69) includes $\alpha = 0.54$, $X_0 = 11.8$ g cm$^{-2}$, and $R_0 = 4.9$ GV/c, with a source spectrum rigidity index of $-2.35$. In principle, all other secondary-to-primary ratios should be consistent with the same parameter set. This is generally found to be the case. As a state-of-the-art application of the weighted-slab technique from Reference 102, we again refer to Reference 69. This is applicable to stable nuclei only, but includes energy losses and gains subject to the limitations described in Section 2.8. They apply the method to one-dimensional disk-halo diffusion, convection, turbulent diffusion, and reacceleration models cast in weighted-slab form. Figure 8 shows B/C and (Sc+Ti+V)/Fe, referred to as sub-Fe/Fe in Reference 69. Clearly, the models cannot be distinguished on the basis of these types of data alone, and they all provide an adequate fit. This shows the

Figure 8

(a) Boron-to-carbon ratio (B/C) and (b) sub-Fe/Fe data compilation, compared with four models treated by the modified weighted-slab technique. $\Phi$ is the cosmic-ray modulation potential in MV. Figure adapted from Reference 69 with permission from the American Astronomical Society.
The figures depict the cosmic-ray propagation in the Galaxy, showing the energy dependence of B/C and (Sc+Ti+V)/Fe ratios.

**Figure a**
- **Standard diffusion model**
- **Turbulent diffusion**
- **Wind model**
- **Minimal reacceleration**
- **ACE-98**
- **ISEE-3**
- **Voyager 1, 2**
- **HEAO-3**
- **UNH (CRID) 1977 (Chappell, Webber)**
- **University of Chicago 1973–74 (Dwyer et al.)**
- **Webber et al. 1985**

**Figure b**
- **Standard diffusion model**
- **Turbulent diffusion**
- **Wind model**
- **Minimal reacceleration**
- **HEAO-3 (Binns et al.)**
- **HEAO-3-C2**
- **Voyager 1, 2**
- **Ichimura & Shibata**
- **ISEE-3**
- **University of Chicago 1973–74 (Dwyer et al.)**
- **IMP-8**
importance of using other species as well as the ones used here. The models can be used to obtain the injection spectrum of primaries, and they find an index of 2.3–2.4 for C and Fe in the energy range 0.5–100 TeV, with the propagated spectrum and data shown in Figure 9.

It has been claimed that no break in \( X(R) \) is required to fit Voyager 2 outer-heliosphere B/C, N/O, and sub-Fe/Fe data extended to 1.5 GeV, plus High Energy Astrophysical Observatory-3 (HEAO-3) data at higher energies, data, and adopting suitable modulation levels (123). Voyager 2 provides a unique data set because of the lower solar modulation in the outer heliosphere.

We consider now explicit models in the sense of Section 2.9. The same procedure is adopted, with \( D_{\infty}(R) \) replacing \( X(R) \). Again, an ad hoc break in \( D_{\infty}(R) \) is required in the absence of other mechanisms. Because of the unphysical nature of such a \( D_{\infty}(R) \), there have been many attempts to find a better explanation, including convection, reacceleration/wave damping, and local sources.

Convection implies an energy-independent escape from the Galaxy so that it dominates at low energies as the diffusion rate decreases, reducing to a simple low-energy asymptotic \( X(R) \propto \beta \). However, this does not resemble the observed B/C energy dependence, as it is too monotonic (67). Furthermore, quite severe limits on the convection velocity come from unstable nuclei.

Reacceleration affects the secondary-to-primary ratio, as described in Section 2.5. Many papers have shown how B/C and other ratios can be reproduced with reacceleration at a plausible level, with no ad hoc break in the diffusion coefficient. An example is shown in Figure 8. In Reference 94 this model is applied to recent ACE (Li, Be, B, C) data. Because reacceleration must be present at some level if diffusion occurs on moving scatterers (e.g., Alfvén waves), this mechanism is favored but not proven. Direct evidence for reacceleration could come from certain K-capture nuclei (Section 3.4). Note that reacceleration requires a smaller value of \( \alpha \), typically 0.3–0.4, consistent with Kolmogorov turbulence, which helps solve the problems with anisotropy (Section 3.5).

Closely related to reacceleration is wave damping, as described in Section 2.5. This can satisfactorily reproduce B/C, protons, and antiprotons, as shown in Figure 10, and also other data (75). The result of this process is a very sharp rise of the diffusion coefficient at rigidities less than approximately 1.5 GV. The Kolmogorov-type dependence is not very successful in this scheme, whereas a Kraichnan-type dependence works better with a high-rigidity asymptotic \( D \sim R^{0.5} \) (panel a in Figure 10).

Quite a different set of parameters has been proposed (68): \( \alpha = 0.7–0.9 \) and injection index \( \approx 2.0 \), on the basis of fitting many species simultaneously, and it is suggested to produce the B/C low-energy decrease from convection. Such a large \( \alpha \) would give problems for the anisotropy (Section 3.5). A related analysis (110) claims

---

**Figure 9**

Data compilation for spectra of (a) C and (b) Fe, compared with four models treated by the modified weighted-slab technique. \( \Phi \) is the cosmic-ray modulation potential in MV. Figure adapted from Reference 69 with permission from the American Astronomical Society.
to exclude the Kolmogorov-type spectrum. An alternative explanation for the fall off in B/C at low energies invokes a weakly nonlinear (in contrast to quasi-linear) transport theory of CRs in turbulent galactic magnetic fields (124, 125).

Another, simpler, explanation of the B/C energy dependence is the local-source model (104, 126) in which part of the primary CR has an additional local component.
Because the secondary flux must come from the Galaxy at large (the local secondaries being negligible), a steep local primary source will cause $B/C$ to decrease at low energies. The known existence of the Local Bubble containing the Sun, and its probable origin in a few supernovae in the last few million years, makes this plausible, but hard to prove. However, it might be possible if CR composition at low energies were found to have anomalies, indicating a younger age compared to high-energy CR. Davis et al. (104) claim that if $B/C$ is fitted in such a model, then sub-$Fe/Fe$ cannot be fitted by the same model. However, an acceptable fit to this and other data is found in Reference 126 using a diffusion model for the large-scale component.

3.2. Unstable Secondary-to-Primary Ratios: Radioactive Clocks

The five unstable secondary nuclei that live long enough to be useful probes of CR propagation are $^{14}C$, $^{10}Be$, $^{26}Al$, $^{36}Cl$, and $^{54}Mn$, with properties summarized in References 101, 126, and 127. $^{10}Be$ is the longest lived and best measured. The theory is presented in Section 2.2. On the basis of these isotopes and updated cross sections (128), the halo height $z_h = 4–6$ kpc, consistent with earlier estimates of 3–7 kpc (98) and 4–12 kpc (67). Figure 11 compares $^{10}Be/^{9}Be$ with models, where the ISOMAX $^{10}Be$ measurements (129) up to 2 GeV (and hence longer decay lifetime) are consistent with the fit to the other data, although the statistics are not very constraining.

The data are often interpreted in terms of the leaky-box model, but this is misleading (108, 127, 131). For the formulae and the detailed procedure for the leaky-box model interpretation, see Reference 132. Luckily, the leaky-box-model surviving fraction can be converted to physically meaningful quantities (131) for a given model. For example, in a simple diffusive halo model, the surviving fraction determines the diffusion coefficient, which can be combined with stable secondary-to-primary ratios to derive the halo size. Typical results are $D_{xx} = (3 - 5) \times 10^{28}$ cm$^2$ s$^{-1}$ (at 3 GV) and $z_h = 4$ kpc. We can then compare the leaky-box model's escape time of $\approx 10^7$ yr with the actual time for CRs to reach the halo boundary after leaving their sources, the
latter being typically an order of magnitude larger. The leaky-box model gas density is typically $0.3 \text{ cm}^{-3}$, compared to the actual average density of $0.03 \text{ cm}^{-3}$ for a 4-kpc halo height, again an order-of-magnitude difference.

Because radioactive secondaries travel only a few hundred parsecs before decaying, it is sometimes believed (133) that they cannot give information on the propagation region of CRs. This is somewhat misleading because it is precisely the combination of stable and radioactive data that allows this. The radioactives determine the diffusion coefficient, which then allows the size of the full propagation region to be determined from the stable secondary-to-primary ratio (where the CRs do come from the entire containment region). This of course assumes the diffusion coefficient does not vary greatly from the local region to the full volume.

The effect of the Local Bubble surrounding the Sun on the interpretation of radioactive nuclei was pointed out in Reference 127. The flux of unstable secondaries is reduced because they are not produced in a gas-depleted region around the Sun. Because they will decay before reaching us, this could lead to an overestimate of the halo size if interpreted in a simple diffusive halo model. This effect would be reduced if the diffusion coefficient in the local region were larger than the large-scale value. In fact, the situation is even more complex. According to References 86, 87, and 134, the Sun left the Local Bubble approximately $10^5$ years ago after spending several million years inside, and we now live in the collection of local interstellar clouds, with HI density approximately $0.2 \text{ cm}^{-3}$ and 35 pc extent. This aspect of the problem for CR propagation has not yet been addressed.

### 3.3. K-Capture Isotopes and Acceleration Delay

Three isotopes produced in explosive nucleosynthesis, $^{59}\text{Ni}$ ($7.6 \times 10^4 \text{ y}$), $^{57}\text{Co}$ (0.74 y), $^{56}\text{Ni}$ (6 d), decay only by K-capture. If acceleration occurs before decay, the decay will be suppressed because the nuclei are stripped. $^{56}\text{Ni}$ is absent as expected, but the other two nuclei are more interesting. Wiedenbeck et al. (135) used ACE data on these nuclei to show that the delay between synthesis and acceleration is long compared with the $^{59}\text{Ni}$ decay time, unless significant $^{59}\text{Co}$ is synthesized in supernovae. Considering theoretical $^{59}\text{Co}$ yields, they conclude on a delay of $\geq 10^5$ years. This is inconsistent with models in which supernovae accelerate their own ejecta, but consistent with acceleration of existing interstellar material. However, the possibility of in-flight electron attachment complicates the analysis (135). For further discussion, see Reference 133. A result from TIGER on Co/Ni at 1–5 GeV/n (136) supports the acceleration delay at higher energies as well.

### 3.4. K-Capture Isotopes and Reacceleration

There have been analyses using ACE data for $^{49}\text{V}$, $^{51}\text{V}$, $^{51}\text{Cr}$, $^{49}\text{Ti}$, and other nuclei (73, 74, 137). Niebur et al. (137) used ACE data on $^{47}\text{Ar}$, $^{44}\text{Ti}$, $^{49}\text{V}$, $^{51}\text{Cr}$, $^{55}\text{Fe}$, and $^{57}\text{Co}$, with inconclusive results. They found that although $^{51}\text{V}/^{51}\text{Cr}$ was in better agreement with reacceleration models, $^{49}\text{Ti}$ gave the opposite result. Jones et al. (74) found that although V/Cr ratios are in slightly better agreement with models
that include reacceleration, ratios involving Ti were inconclusive. The main problem is the accuracy of the fragmentation cross sections (126). Discussion of ACE and previous results can be found in section 3 of Reference 133.

3.5. Anisotropy

High isotropy is a distinctive quality of CRs observed on Earth. The global leakage of CRs from the Galaxy and the contribution of individual sources lead to anisotropy, but the trajectories of energetic charged particles are highly tangled by regular and stochastic interstellar magnetic fields that isotropize the CR angular distribution. This makes the direct association of detected CR particles, except for the highest-energy particles, with their sources difficult or even impossible. Observations give the amplitude of the first angular harmonic of anisotropy at the level of \( \delta \sim 10^{-3} \) in the energy range \( 10^{12} \) to \( 10^{14} \) eV, where the most reliable data are available (Figure 12) (see References 138 and 139). The angular distribution of particles at lower energies is significantly modulated by the solar wind. The statistics at higher energies are not good enough yet, but the measurements indicate the anisotropy amplitude at a level of a few percent at \( 10^{16} - 10^{18} \) eV. The data of the Super-Kamiokande-I detector (140) allowed accurate two-dimensional mapping of CR anisotropy at \( 10^{13} \) eV [see also the Tibet Air-Shower Array results (138)]. After correction for atmospheric effects, the deviation from the isotropic event rate is 0.1\%, with a statistical significance > 5\( \sigma \) and direction to maximum excess at roughly \( \alpha = 75^\circ \) and \( \delta = -5^\circ \).

The amplitude of anisotropy in the diffusion approximation is given by the following equation, which includes pure diffusion and convection terms:

\[
\delta = -\left[ 3D \nabla f + \left( \frac{\partial f}{\partial p} \right) v_p \right] / \langle f \rangle,
\]

where \( f \) is defined after Equation 1 (see Reference 8). Here, \( D \) is the
diffusion tensor, and it is assumed that the magnetic inhomogeneities that scatter CR particles are frozen in the background medium, moving with velocity \( u \ll v \), which gives rise to the convection term (also known as the Compton-Getting term).

The Compton-Getting anisotropy is equal to \((\gamma + 2)u/c\) for ultrarelativistic CRs with a power-law spectrum \( I(E) \sim p^2 f(p) \sim E^{-\gamma} \). The motion of the solar system through the local ISM produces the constant term in the energy dependence of the anisotropy \( \sim 4 \times 10^{-4} \), with a maximum intensity in the general direction of the galactic center region, which does not agree with the data at \( 10^{12} \)–\( 10^{14} \) eV, which point to an excess intensity from the anticenter hemisphere. The convection effect is outweighed by the diffusion anisotropy, which is due to the nonuniform distribution of CR in the Galaxy. The systematic decrease of CR intensity to the periphery of the Galaxy and the CR fluctuations produced by nearby SNRs are almost equally important for the formation of the local CR gradient.

Calculations of CR anisotropy (142) are illustrated in Figure 12, where the effect of global leakage from the Galaxy and the overall contribution of known SNRs with distances up to 1 kpc are shown separately. Two basic versions of the flat-halo diffusion model—the plain diffusion model with \( D \sim E^{0.54} \) and the reacceleration model with \( D_{xx} \sim E^{3} \)—were used in the calculations. The values of \( D_{xx} \) were taken from Reference 69. The diffusion from a few nearby SNRs with different ages and distances results in a nonmonotonic dependence of anisotropy on energy. The flux from the Vela SNR probably dominates over other SNRs at energies below approximately \( 6 \times 10^{13} \) eV in the plain diffusion model and below approximately \( 10^{14} \) eV for the reacceleration model. The results of these calculations indicate that the diffusion model with reacceleration is compatible with the data on CR anisotropy within a factor of approximately three. However, it seems that the plain diffusion model, with its relatively strong dependence of diffusion on energy, predicts a too-large anisotropy at \( E > 10^{14} \) eV. Note that CR diffusion here is assumed to be isotropic, which is a seriously simplifying assumption, but difficult to avoid because of the complicated and unknown detailed structure of the galactic magnetic field. The presence of a large-scale random magnetic field justifies the approximation of isotropic diffusion on scales larger than a few hundred parsecs, but the anisotropy is a characteristic that is very sensitive to the local surroundings of the solar system, including the direction of magnetic field and the value of the diffusion tensor (see discussion in Reference 8).

### 3.6. Diffuse Galactic Gamma Rays

Gamma rays (above approximately 100 MeV) from the ISM hold great promise for CR studies because we can observe them throughout the Galaxy, not just from the local region of direct CR measurements. The complementarity of \( \gamma \)-rays and direct CR measurements can be exploited to learn about CR origin and propagation. Gamma rays are produced in the ISM by interactions of CR protons, He \((\pi^0\text{-decay})\), and electrons (bremsstrahlung) with gas, and electrons with the ISRF via inverse-Compton scattering. For details of the processes, the reader is referred to References 39, 81, 88, and 113. Additional astronomical material comes into play, such as the
distribution of atomic and molecular gas, and the ISRF. In fact, γ-rays provide an important independent handle on molecular hydrogen and its relation to its CO molecular tracer, which has taken its place beside more traditional determinations.

Historically, observations started with the OSO-III satellite in 1968, followed by the SAS-2 (1972), the COS-B (1975–1982), and the Compton Gamma-Ray Observatory (CGRO) (1991–2000). Each of these experiments represented a significant leap forward with respect to its predecessor. SAS-2 established the existence of emission from the ISM and allowed a first attempt to derive the CR distribution and also constrain the CR halo size (143). With COS-B, the CR distribution could be better derived and it was found not to follow the canonical distribution of SNRs, which posed a problem for the SNR origin of CRs. A relatively dependable value for the CO-H$_2$ relation was derived. With the Energetic Gamma-Ray Experiment Telescope (EGRET) and COMPTEL on the CGRO, the improvement in data quality was sufficient to allow such studies to be performed in much greater detail. There is now so much relevant information and theory that a rather realistic approach is justified and indeed necessary. At this point, the best approach seems to be explicit modeling of the high-energy Galaxy, putting in concepts from CR sources and propagation, galactic structure, and so on. The idea of a single model to reproduce both CR and γ-ray (and other) data simultaneously arose naturally and is the goal of the GALPROP project (see Section 2.10). For a recent review, see Reference 144.

To illustrate the current state of the art, we show spectra and profiles from a recent GALPROP model compared with CGRO/EGRET and COMPTEL data. These are based on References 113 and 145, where full details can be found. The first model is based simply on the directly measured CR spectra together with a radial gradient in the CR sources. It is immediately clear that the spectrum is not well predicted, being below the EGRET data for energies above 1 GeV. However, remembering that this is an unfitted prediction, it does show that the basic assumption that γ-rays are produced in CR interactions is correct. The factor of two difference tells us something about the remaining uncertainties. The extent to which the CR spectra have to be modified to get a good fit that includes the excess in the GeV region is shown in Figure 13, and the resulting γ-ray-model prediction in Figure 14. The difference between the directly observed and modified spectra is within plausible limits, considering solar modulation and spatial fluctuations in the Galaxy on scales >100 pc. The model has been tuned to fit the CR secondary antiprotons also produced in proton-proton interactions, on the assumption that these antiprotons have smaller spatial fluctuations than protons in the Galaxy. A detailed justification is lacking, however, so this so-called optimized model is just an existence proof rather than a conclusive result.

Other more drastic modifications to the CR spectrum have been proposed (78) as follows:

- A very hard electron injection spectrum, which could be possible invoking large fluctuations due to energy losses and the stochastic nature of SNRs in space and time: The solar vicinity is not necessarily a typical place for electron measurements to be representative (81, 146). However, the variations required are even larger than can reasonably be expected (113), so this model seems unlikely.
A hard proton spectrum, again invoking spatial variations in the Galaxy so the solar vicinity is not typical: This is more difficult than for electrons because the proton energy losses are negligible. It turns out this possibility can be ruled out on the basis of antiproton measurements. Too many antiprotons would be produced by the same protons that generate the γ-rays (78, 113, 147). A related suggestion invokes the dispersion in the radio spectral indices of SNRs, which indicates a dispersion in electron indices and, if assumed to apply to CR protons (148), could produce the GeV excess. This should also be tested against antiprotons.

It could be that a completely different source of the excess GeV γ-rays is present, and the possibility of a dark matter origin was pursued (150) but found to produce an excess of CR antiprotons (151). We will not enter this debate here because it is not related directly to the problem of CR propagation. It does, however, show how a good understanding of the CR-induced γ-rays in the Galaxy is essential to the study of potentially more fundamental physics (152).

The angular distribution of γ-rays provides an essential test for any model. The problem with the large expected gradient from SNRs is critical. The distribution of SNRs is hard to measure because of selection effects, so this problem could be safely ignored in the past. However, now the distribution of pulsars can be determined with reasonable accuracy, and this should trace SNRs because supernovae are pulsar progenitors. The pulsar gradient is indeed larger than originally deduced from γ-rays, as shown in Figure 15. The distribution of SNRs in external galaxies shows a similar
Figure 14
Gamma-ray spectrum of the inner Galaxy (330° < l < 30°, |b| < 5°) for models based on the directly observed cosmic-ray spectra and modified spectra shown in Figure 13. l is galactic longitude; b is galactic latitude. This is an update of the spectra shown in Reference 113.

Figure 15
Possible cosmic-ray (CR) source distributions as a function of a Galactocentric radius R. Pulsars (blue line) and supernova remnants (vertical bars) as deduced from γ-rays assuming a constant H_2-to-CO relation (red line) (145).
Gradient to pulsars (153). The problem therefore remains, but a possible solution may be found in another uncertainty—the galactic distribution of molecular hydrogen. Because we must rely on the CO tracer molecule, any variation in the relation of CO to H$_2$ will affect the interpretation of the $\gamma$-ray data. Strong et al. (145) noted that there is independent evidence for an increase in the ratio of H$_2$ to CO with a Galactocentric radius, related to galactic metallicity gradients, and that this can resolve the problem, allowing CR sources to follow the SNR distribution as traced by pulsars. Figure 16 shows longitude and latitude profiles based on such a model, showing a satisfactory agreement with EGRET data. However, the magnitude of the variation in the ratio of H$_2$ to CO, or its relation to metallicity and other effects, is uncertain, so the issue will need to be studied more. Future data from GLAST will help by giving much finer angular resolution and the possibility of better separating the molecular and atomic hydrogen components. Breitschwerdt et al. (70) proposed another possible solution to the gradient puzzle in terms of a radially dependent galactic wind (see Sections 2.4 and 2.12).

Note the large contribution from the inverse-Compton emission to the spectra and profiles (Figure 14 and Figure 16). This is the reason why the predicted emission is in good agreement with the EGRET data right up to the highest latitudes (where the
3.7. Antiprotons and Positrons

The spectrum and origin of antiprotons in CRs has been a matter of active debate since the first reported detections in balloon flights (154, 155). There is a consensus that most of the CR antiprotons observed near Earth are secondaries (156). Owing to the kinematics of secondary production, the spectrum of antiprotons has a unique shape that distinguishes it from other CR species. It peaks at approximately 2 GeV, decreasing sharply toward lower energies. In addition to secondary antiprotons, there are possible sources of primary antiprotons. Those most often discussed are dark matter particle annihilation and evaporation of primordial black holes.

In recent years, new data with large statistics on both low- and high-energy antiproton fluxes have become available (157–163), thanks mostly to continuous improvements of the BESS instrument and its systematic launches every one to two years. This allows one to test models of CR propagation and heliospheric modulation. Additionally, accurate calculation of the secondary antiproton flux predicts the background for searches for exotic signals such as weakly interacting massive particle (WIMP) annihilation.

Despite numerous efforts and overall agreement on the secondary nature of the majority of CR antiprotons, published estimates of the expected flux differ significantly (see, e.g., figure 3 in Reference 159). Calculation of the secondary antiproton flux is a complicated task. The major sources of uncertainties are threefold: (a) incomplete knowledge of cross sections for antiproton production, annihilation, and scattering; (b) parameters and models of particle propagation in the Galaxy; and (c) modulation in the heliosphere. Although the interstellar antiproton flux is affected only by uncertainties in the cross sections and propagation models, the final comparison with experiment can be made only after correcting for the solar modulation. Besides, the spectra of CR nucleons have been directly measured only inside the heliosphere, whereas we need to know the spectrum outside, in interstellar space, to correctly compute the antiproton production rate.

Moskalenko et al. (117) recently showed that accurate antiproton measurements during the last solar minimum (1995–1997) (159) are inconsistent with existing propagation models at the ~40% level at approximately 2 GeV, while the stated measurement uncertainties in this energy range are now ~20%. The conventional models based on (a) local CR measurements, (b) simple energy dependence of the diffusion coefficient, and (c) uniform CR source spectra throughout the Galaxy fail to reproduce simultaneously both the secondary-to-primary nuclei ratio and antiproton flux.
A reacceleration model designed to match secondary-to-primary nuclei ratios produces too few antiprotons because the diffusion coefficient is too large. Models without reacceleration can reproduce the antiproton flux, but cannot explain the low-energy decrease in the secondary-to-primary nuclei ratios. To be consistent with both, the introduction of breaks in the diffusion coefficient and the injection spectrum is required, which would suggest new phenomena in particle acceleration and propagation. A solution in terms of propagation models requires a break in the diffusion coefficient at a few GeV, which has been interpreted as a change in the propagation mode (117). This calculation employs a modern steady-state drift model of propagation in the heliosphere to predict the proton and antiproton fluxes near Earth for different modulation levels and magnetic polarity.

If our local environment influences the spectrum of CRs, then it is possible to solve the problem by invoking a fresh, unprocessed nuclei component at low energies (126), which may be produced in the Local Bubble. The idea is that primary CRs such as C and O have a local low-energy component, whereas secondary CRs such as B are produced Galaxy wide over the confinement time of 10–100 Myr. In this way, an excess of B, which appears when propagation parameters are tuned to match the $\bar{p}$ data, can be eliminated by an additional local source of C (see Section 3.1). The model appears to describe a variety of CR data, but at the cost of additional parameters.

A consistent $\bar{p}$ flux can be obtained in reacceleration models if there are additional sources of protons $\lesssim20$ GeV (164). This energy is above the $\bar{p}$ production threshold and effectively produces $\bar{p}s$ at $\lesssim2$ GeV. The intensity and spectral shape of this component can be derived by combining restrictions from $\bar{p}s$ and diffuse $\gamma$-rays.

Clearly, accurate measurements of $\bar{p}$ flux are the key to testing current propagation models. The GeV excess in diffuse $\gamma$-rays discussed in the sections above may be interpreted in terms of fluctuations in the CR density in the Galaxy, and if so, the antiproton and $\gamma$-ray excesses have the same origin (113). If new measurements confirm the $\bar{p}$ excess, current propagation and/or modulation models will be challenged. If not, it will be evidence that the reacceleration model is currently the best one to describe the data. Here we must wait for the next BESS-Polar flight, which should bring much larger statistics, and for Pamela (165), currently in orbit.

Positrons in CRs were discovered in 1964 (166). The ratio of positrons to electrons is approximately 0.1 above $\sim1$ GeV, as measured near Earth. Secondary positrons in CRs are mostly the decay products of charged pions and kaons ($\pi^+, K^+$) produced in CR interactions with interstellar gas. In general, the calculation of the positron production by CRs agrees with the data and indicates that the majority of CR positrons are secondary (39, 167). Some small fraction of positrons may also come from sources (168) such as pulsar winds (169) or WIMP annihilations (170).

In interstellar space, the secondary positron flux below $\sim1$ GeV is comparable to the electron flux, which makes positrons non-negligible contributors to the diffuse $\gamma$-ray emission in the MeV range (113). The contribution of secondary positrons and electrons in diffuse galactic $\gamma$-ray emission models essentially improves the agreement with the data, making secondary positrons effectively detectable in $\gamma$-rays.

The accuracy of spacecraft and balloon-borne experiments has increased substantially during the past decade. The data taken by various instruments, including AMS,
CAPRICE, and HEAT (41, 171–173), are in agreement with each other within the error bars, which are, however, still quite large. A recent measurement of CR positron fraction by the HEAT instrument (42) indicates an excess (feature) between 5 and 7 GeV at the level of $\sim 3\sigma$ over the smooth prediction of a propagation model (39). When all the HEAT flights are combined, the data show an excess above $\sim 8$ GeV, with the most significant point at $\sim 8$ GeV (42). The Pamela satellite (165) currently in orbit will have very good positron statistics.

### 3.8. Electrons and Synchrotron Radiation

CR electrons require a separate treatment from nuclei because of their rapid energy losses and their link with synchrotron radiation. Their low energy density compared with nuclei ($\approx 1\%$) is not yet understood. Standard SNR shock acceleration does not predict such a ratio, and it is normally a free parameter in models (e.g., Reference 174). Direct measurements extend from MeV to TeV. At low energies, solar modulation is so large that the interstellar fluxes are completely unknown. At TeV energies, the statistics are very poor, but new experiments are in progress. Young SNRs contain TeV electrons, as shown by their nonthermal X-ray emission, and energies up to 100 TeV are possible (175). The rapid energy losses mean that contributions from old nearby SNRs such as Loop I, Vela, the Cygnus Loop, and MonoGem could make an important contribution to the local electron spectrum above 100 GeV and dominate above 1 TeV (see Figure 4) (37). The rapid energy losses of electrons also produce time- and space-dependent effects, as described in the next section.

The essential propagation effects are seen in Figure 13. At energies below 1 GeV, the interstellar spectrum is flatter than the injection spectrum owing to Coulomb losses. There is an intermediate energy range where the spectrum is similar to that at injection, and at high energies it steepens owing to inverse Compton and synchrotron losses and energy-dependent diffusion. For propagation models, we need an injection spectrum that reproduces the observations. A typical spectral index is $2.4$, similar to nuclei, which produces the observed high-energy slope of $3.3$ (113). The synchrotron spectral index is $\beta_{\nu} = (\gamma - 1)/2$ for an electron index $\gamma$. The observed $\beta_{e} = 0.6 - 1$, increasing with frequency, implies $\gamma = 2.4 - 3$, increasing with energy, in satisfactory agreement. Detailed geometrical models of the galactic synchrotron emission have been constructed (176, 177). A more physical interpretation requires propagation and magnetic field modeling, as, for example, in Reference 81. Combined with $\gamma$-ray data, this allows the magnetic field to be determined as a function of a Galactocentric radius independent of other techniques (81). Since the work reported in Reference 81, a great deal of new radio data have become available both from new ground-based surveys and from balloons and satellites for cosmology (e.g., WMAP). Because galactic synchrotron is an essential foreground for cosmic microwave background studies, there is a large overlap of interest between these fields. Exploitation of these data for CR studies has only just begun, but is expected to accelerate with the forthcoming Planck mission.

Radio continuum observations of other galaxies provide a complementary view of electrons via synchrotron radiation (178, 179). The classical case is the edge-on
galaxy NGC891 (180, 181), which has a nonthermal halo extending to several kiloparsecs. This helped give credence to the idea of a large halo in our Galaxy. More recent observations of this and other edge-on galaxies confirm large halos (182). The spectral index variations due to energy losses provide a test of propagation in principle, although this has not been very fruitful up until now. This effect can also be measured in starburst galaxies (183). Face-on galaxies such as M51 show marked spiral structure in the continuum (178), but this is probably mainly a magnetic field effect. It is not possible to separate CR and magnetic field variations reliably, although this is often attempted assuming equipartition of energy between CRs and field, which although having no firmly established physical basis, nevertheless gives field values quite similar to other methods (179, 184).

The tight radio-continuum/far-infrared correlation for galaxies (185–187) is worth mentioning here. It holds both within galaxies (down to 100 pc or less) and from galaxy to galaxy, over several orders of magnitude in luminosity. It presumably contains clues to CR origin and propagation. The CR calorimeter (188) is the simplest interpretation. Simply relating both CR production and ultraviolet (reprocessed to far-infrared) to star formation rates is apparently insufficient to explain the close correlation, and this has given rise to several different interpretations, including hydrostatic regulation, without being very conclusive (185–187).

### 3.9. Time-Dependent and Space-Dependent Effects

CR propagation is usually modeled as a spatially smooth, temporally steady-state process. Because of the rapid diffusion and long containment timescales in the Galaxy, this is usually a sufficient approximation. However, there are cases where it breaks down. The rapid energy loss of electrons above 100 GeV and the stochastic nature of their sources produce spatial and temporal variations. Supernovae are stochastic events and each SNR source of CRs lasts only perhaps $10^4 - 10^5$ years, so that they leave their imprint on the distribution of electrons. This leads to large fluctuations in the CR electron density at high energies, so that the electron spectrum measured near the Sun may not be typical (113). A statistical calculation of this effect can be found in References 146 and 189, and for local SNRs in Reference 37. These effects are much smaller for nucleons because in this case there are essentially no energy losses, but they can still be important (190). For the theory of CR fluctuations for a Galaxy with random SNR events, see Reference 142 and Section 3.5. Such effects can influence the B/C ratio (111, 116) mainly through variations in the primary spectra. The Local Bubble can also have an effect on the energy dependence of B/C (126). Dispersion in CR injection spectra from SNRs may cause the locally observed spectrum to deviate from the average (148).

### SUMMARY POINTS

1. Presently, progress in CR research, with the new generation of detectors for both direct and indirect measurements, is rapid. However, it is still hard to
pin down particular theories. Even now the origin of the nucleonic component is not proven (34), although SNRs are the leading candidates.

2. Considering all relevant data, both direct (particles) and indirect (γ-ray, synchrotron) measurements, is important.

3. Increase in computing power has made many of the old approximations for interpreting CR data unnecessary.

4. New high-quality data will require detailed numerical models.

FUTURE ISSUES
We end by listing some of the open questions regarding CR propagation, which may be answered with observations in the near future.

1. The interpretation of the energy-dependence of the secondary-to-primary ratio, which requires accurate measurements at both low and high energies, is an open question.

2. The size of the propagation region—the existence of an extended halo—requires measurements of radioactive species over a broad energy range.

3. What are the relative roles of diffusion, convection, and reacceleration?

4. What is the importance of local sources to the primary CR flux?

5. What is the origin of the GeV excess in γ-rays relative to the prediction based on locally observed CRs?

6. Is the CR source distribution similar to SNRs?

7. Are positrons and antiprotons fully explained as secondaries from primary CRs, or is there a perhaps exotic excess?

8. What is the relation of CR-dynamical models of the Galaxy to CR propagation theory?

DISCLOSURE STATEMENT
The authors are not aware of any biases that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS
I.V.M. acknowledges partial support from a NASA Astronomy and Physics Research and Analysis program grant.
LITERATURE CITED


Related Resources

GALPROP Web site: [http://galprop.stanford.edu](http://galprop.stanford.edu)
GLAST Web site: [http://glast.stanford.edu](http://glast.stanford.edu)