Chapter 10

Conclusion

Two decades have elapsed since the second edition of *Dust in the Galactic Environment* was completed in 2002. I concluded its final chapter with a discussion of open questions, and signed off by remarking that if I should ever be foolhardy enough to attempt a third edition perhaps some of the answers would adorn its pages. These things have come to pass. Of course, we have not answered all of the questions; but a great deal of progress has been made, and we have a clearer vision of what questions we should be addressing next. This chapter summarizes what we have learned and what we might hope to learn in the future.

Before getting into specifics, however, it is interesting to consider how the field itself has changed. In my experience, astronomical research has evolved to become progressively more interdisciplinary over the past half century. This may reflect a general trend in all science, of course, but in astronomy it seems to have arisen, at least in part, from a few game-changing discoveries. A rapid growth in the inventory of known interstellar molecules demonstrated that space environments host chemical as well as physical complexities. Subsequent identification of presolar grains by the application of geochemical methods to studies of meteorites opened new windows, not only on the nature and origin of cosmic dust but also on the nuclear processes that created their elements inside stars. Next came the discovery of exoplanets, including potential Earth-analogs, detected by well-honed astronomical techniques but requiring expertize in many disciplines—geosciences, environmental sciences, biosciences—to attempt any realistic assessment of their characteristics. Simultaneously, our understanding of the processes that forged our own planetary system are being clarified by research programs that range from analyses of extinct radionuclides in meteorites to cometary and asteroidal rendezvous missions and high-resolution spatial and spectral imaging of protoplanetary disks around potential early solar analogs.

Research on dust is emblematic of this trend because it is pertinent to all aspects of the science. A useful metric is the number of relevant review papers. Citations in

doi:10.1088/2514-3433/ac7204ch10

this text published in review journals such as the *Annual Reviews* series are not limited to the those specific to astronomy. Excluding conference proceedings, the total number of cited reviews is 59, of which 42 have appeared in the last 20 years. Of course, by no means all of these reviews are specifically "about" dust, but all were deemed sufficiently relevant for citation (the count was made retrospectively!). I doubt that this statistic implies any general increase in the number of review papers being written over the past two decades, but instead a growing synergy between the astrophysics of dust and other areas of research.

10.1 An Overview of Dust Populations

A wide range of observational and empirical constraints on the properties of cosmic dust are described in the preceding chapters. The data imply a need for models that combine different grain populations to account for different aspects of the data: the prime suspects are summarized in Table 10.1. These populations are distinguished by their spectral properties and inferred differences in composition, size distribution and physical structure (amorphous or crystalline, mantled or unmantled, monolithic or composite, rounded or oblate). A unified model for interstellar dust must be capable of matching all known constraints, invoking substances and size distributions that are physically reasonable and have a probable source in terms of stellar mass-loss processes and/or production in the ISM itself. A *unique* model would accomplish all of this without ambiguity, but this is not yet achievable because the available data do not impose absolute constraints (different assumptions can yield similarly good fits), and also perhaps because our models are not yet sufficiently realistic.

Table 10.1. Interstellar Dust Constituents Implied by Observational and Empirical Evidence.

Property	Amorphous Silicates	Mantles	Large Carbon	Bump Grains	Small Carbon
Composition	$MgSiO_3$, etc. with Fe-rich inclusions	ices a-C ORM?	a-C graphite (SiC)	graphite	PAHs diamond?
Origin	Stardust	ISM	Stardust	Both?	Both?
Size (μm)	$a \lesssim 1$	$\Delta a \lesssim 0.1$	$a \lesssim 10$	$a \lesssim 0.01$	$a \lesssim 0.01$
Elements depleted	O, Mg, Si, Fe	O, N, C	C	C	C
Extinction	UV–IR	UV–IR	UV–IR	mid-UV	UV
Alignment?	yes	(yes)	no?	no?	no?
Absorption features	9.7, 19 μm	3 μm, etc.	3.4 μm?	217 nm	3.4 μm?
Emission features	CS only?	—	—	—	PAH bands
Continuum emission	FIR	FIR	FIR	MIR	MIR

Notes: Composite particles may contain mixtures of these constituents. Bump and small carbon grains may belong to the same population. SiC is inferred to be present from studies of meteorites, not by direct observation. Only abundant depleted elements are listed.

The goal of this section is to present a general overview, highlighting areas of both convergence and ambiguity, with a view to future progress toward a robust, unified theory of cosmic dust. No citations of specific models are included here. See Sections 3.6 and 4.2.8 for examples of fits to extinction and polarization data, and Section 4.3 for a review of grain alignment theory. See also Sections 3.2.3 (scattering), 3.4.5 (bump), 4.2.7 (spectropolarimetry), 5.2 (refractory dust), 5.3 (ices), 6.2.2 (PAHs), 7.3 (element depletions), and 8.3 (meteoritic stardust) for discussion of constraints placed on models by specific observables. Table 10.1 summarizes the properties of five proposed constituents of interstellar dust. Successful models for dust in the diffuse ISM require at least three: silicates, small carbon, and some other C-rich repository. The chemical composition of interstellar silicates is not uniquely defined, but available data suggest that amorphous magnesium silicates ($MgSiO_3$, Mg_2SiO_4, $MgFeSiO_4$) may be more common than Fe-rich forms. Fe and its oxides may also be present as embedded subparticles (inclusions) in silicates and/or as an dependent population, and some Fe may also be bonded to carbon. It is plausible to suppose that the small carbon and “bump” grains belong to the same population of nanoparticles; they are listed separately in Table 10.1 because they are invoked to explain different spectral features (IR emission, UV absorption), and the agents responsible may differ in detail. A distinct population of larger carbonaceous particles is essential to all models, to account for the continuum opacity of the dust without violating constraints arising from the spectral features, which require both PAH emitters and bump absorbers to be very small. A point of divergence in extant models is the assumed form of these larger particles: proposals include mantles on silicates, aggregates of mixed composition, and an independent population. Which of these is the most realistic choice is a persistent and perplexing question.

The answer to this question is linked to another: where and how does the dust originate? If stellar outflows are major sources (Section 8.4), their contributions divide naturally into distinct O-rich and C-rich populations. A gradual blending of these populations by grain–grain aggregation in clouds (Section 9.1.1) is physically reasonable and indeed expected. The growth and subsequent carbonization of icy mantles on silicate cores as they cycle in and out of dense clouds may also lead to grains of mixed carbon/silicate composition. A diagnostic test of this trend is provided by spectropolarimetry. The observed wavelength-dependence of polarization is best explained if silicate grains are its primary source. Silicates and (where present) ice mantles exhibit polarization peaks corresponding to their absorption features, clearly demonstrating that such particles produce dichroic extinction and respond to alignment. In contrast, C-rich dust generally lacks polarimetric signatures. Both the 217 nm graphite bump and the 3.4 μm hydrocarbon feature have been scrutinized, with almost entirely negative results (two lines of sight with weak bump excesses are the only known exceptions). The smallest grains tend to be least responsive to alignment, so the general lack of bump polarization is not unexpected. Perhaps the 3.4 μm feature also arises primarily in small grains? Perhaps aggregate particles tend to be more rounded than monolithic silicate grains? As it stands, the available evidence is consistent with distinct populations of silicate and carbon grains, of which only the former is a significant agent for polarization.

The apparent need for a distinct silicate population is consistent with the view that most interstellar silicates originate in stellar atmospheres. This setting seems far more conducive to silicate production than the colder, more rarefied environments of interstellar clouds. The argument against this proposition arises from calculations of expected timescales for destruction by shocks, which suggest that silicates are not being replenished rapidly enough by stellar winds to explain their abundance in the ISM; however, more recent reassessments seem to have alleviated this constraint somewhat (see Section 8.4.4 for a detailed discussion). I therefore conclude that stardust is the most likely source of interstellar silicates. In contrast, an interstellar origin for (some) C-rich grain material is very plausible as a product of mantle growth on existing grains. Limits on the abundance of SiC dust (confirmed only in meteorites) suggest that carbon stars are less important as sources of interstellar grains.

A power-law form of the size distribution function $n(a)$ (Equation (3.30)) is typically assumed in (or deduced from) models for the optical properties of the dust, applicable over some appropriate range of a for each component. This mathematical form is physically reasonable as the predicted outcome of grain–grain collisions in both circumstellar and interstellar environments (Sections 8.4.2 and 9.1.1). The models yield acceptable fits to observations of extinction and scattering, with variations attributed to changing conditions between diffuse and dense clouds accommodated by modest differences in the relevant parameters (power-law index, minimum and maximum sizes). The maximum size is often capped at $a_{\max} = 1\ \mu\text{m}$ or less, as much larger grains have minimal effect on the extinction curve over the spectral range typically observed, and too many large grains could exceed abundance constraints, especially for silicates. Nevertheless, there are reasons to challenge this assumption. The most tangible evidence for large ($a > 1\ \mu\text{m}$) grains is found in the presolar meteorite collection (Section 8.3.4), and in the present-day flux of interstellar dust into the solar system (Section 1.3.4). Other phenomena that might best be explained by extending the limit to accommodate larger grains include X-ray scattering halos and core-shine (Section 3.2.3), broadening of ice absorption profiles (Section 5.3.3), and the oxygen depletion anomaly (Section 7.3.3). Refined methods of detecting any *wavelength-independent* component of the total extinction (Section 3.2.5) may be used to further explore this possibility.

Our understanding of the nature and evolution of the ices that accumulate on dust in dense clouds has evolved dramatically in recent decades, primarily as the result of data from infrared space observatories such as ISO and Spitzer, in tandem with laboratory work designed to support these missions. The results give insight into not only the composition and structure of the ices (Section 5.3) but also how they evolve with changing conditions as clouds collapse and begin to form stars (Section 9.3). Successive phases of mantle growth create chemically distinct layers. Subsequently, as they are warmed, the ices undergo selective sublimation, segregation and annealing. Some ice-mantled grains may accrete directly into icy planetesimals in protoplanetary disks, while others sublimate and recondense. All such volatiles retain isotopic signatures of their low-temperature origins, most markedly enhanced D/H ratios, now detected in reservoirs that range from comets and asteroids to the Earth's ocean water.

10.2 Future Prospects

The progress of science is driven by the availability of data and our ability to comprehend the message it contains. Data acquisition in the astronomical sciences entered an era of exponential growth in the latter part of the 20th century, with the deployment of space telescopes such as IRAS, Hubble and ISO, together with ground-based facilities using the latest technologies. The ingenuity of scientists engaged in the task of interpreting the message (and distinguishing the wood from the trees!) led to advances described in this text. The stream of new data shows no sign of abating. At the time of this writing, the James Webb Space Telescope (JWST) is undergoing final *in situ* alignment and testing. Once fully commissioned, it will have a profound influence on virtually all areas of astronomy. I conclude this text with a summary of some key questions relating to our topic that might be answered by future research, some but not all enabled by the JWST.

What follows is a personal selection of important issues, with no claim to be fully comprehensive. I do not address questions relating specifically to dust in distant galaxies, for example, because others are far better qualified to do so. I am certain that a fuller understanding of dust in our home Galaxy and its near neighbors will assist and inform our attempts to assess its broader significance in the cosmos.

10.2.1 The Origin and Structure of Refractory Dust

- Can we better quantify the sources of refractory interstellar dust?
- What is the predominant structure of the grains responsible for continuum opacity? How are elemental C and Fe distributed among these particles?
- How common are the largest (super-micron-sized) grains in the ISM?

These are tantalizing and persistent questions. As previously discussed, stellar atmospheres of appropriate composition, temperature and pressure are proven sources of stardust, and these conditions cannot be replicated in the ISM. This is a problem especially for our understanding of the ecology of interstellar silicates if, as some models suggest, the infusion rate is too low. Estimates focus on mass loss from evolved (AGB) stars; but it should not be overlooked that the circumstellar envelopes of young stars also make silicates, and they also undergo mass loss as protoplanetary disks are cleared out. It may not be possible to detect such particles by isotopic analysis of meteorites: presolar silicates are identified by isotopic signatures specific to evolved stars, and all such fingerprints are homogenized in the condensation zone of a protoplanetary disk, whether it be our own or in another system prior to arrival. Nevertheless, developers of quantitative models for the life cycle of dust should reconsider this possible source.

Carbonaceous dust may originate from stellar or interstellar sources. The nanoparticles (PAH clusters, bump grains) can be understood in terms of collisional fragmentation of either stardust or carbonized grain mantles (probably both). But what of the big grains needed to explain general extinction? If carbonaceous solids are largely synthesized in the ISM, they must presumably originate as mantles on pre-existing dust. In diffuse clouds, elemental C may deplete directly from the gas to grain surfaces; in denser clouds, C-bearing molecules are sequestered in the ices. The evolution of these solids as they cycle though different environments may be traced by infrared spectroscopy, but attempts to do so are limited by the fact that their absorption features are very weak, e.g., in comparison to those of silicates and ices (see Figure 5.5). Thus, very high quality spectra are needed to extract information from (and in some cases even to detect) diagnostic features such as the C–H, C–C, and C=O modes of hydrogenated and organic carbon, and the lattice modes of graphite. The superior sensitivity and resolution enabled by the JWST and its suite of instruments should be transformative. It may be possible, for example, to explore how ice mantles evolve in the outer layers of a dark cloud, where they are subject to photolysis at extinctions near the threshold value. To what extent does the ice crystallize, sublimate or undergo chemical change? Are photolyzed ices converted into organic kerogens prior to complete carbonization? Is the highly aromatic state of carbon dust toward the center of our Galaxy typical of the diffuse ISM in general?

Another interesting problem is the role of iron in the dust. Fe is the most abundant heavy metal. It is highly depleted from the gas and is a major contributor to the total dust mass (e.g., Figure 7.4). Yet there is no clear consensus on what form it takes. Interstellar silicates may account for some, but a stardust origin for silicates favors Mg-rich condensates, with Fe more likely to emerge in metallic or oxidized form. Aggregation might lead to clustering of these products. Certain interplanetary particles (GEMS) that contain Fe-rich inclusions in a silicate structure are suggested as plausible analogs for interstellar silicates. In pure form, Fe lacks spectral signatures in the infrared. Oxides such as FeO produce broad spectral features near 20 μm that blend with the adjacent silicate feature (Figure 8.9), detected in O-rich circumstellar matter but not so far in the ISM.

The smallest grains in the size distribution have been a focus of research for the past several decades, in response to the discovery of mid-infrared line and continuum emission and the inferred presence of transiently-heated nanoparticles. It may be timely now to focus on the largest particles, for reasons outlined in Section 10.1 and touched on elsewhere in this book. Early studies of interstellar extinction and reddening (Sections 1.2.1 and 1.2.2) were motivated by the need to fully account for all sources of attenuation in the visible (including any "neutral" component arising from grains large compared with the wavelength) because of the potential impact on calibration of the cosmic distance scale. Stellar distances measured by parallax are independent of dust properties, but a century ago accurate parallaxes could be measured only for nearby stars ($d \lesssim 10$ pc) that are essentially free of extinction. Modern space astrometry missions, most notably Gaia, have extended this limit by orders of magnitude. Hence it is now possible to use stellar distances to evaluate extinction independent of reddening, effectively inverting the analytical method used in early work. This may result in tighter constraints on the contributions of very large grains to visual extinction in the diffuse ISM; results for dense clouds may provide a new measure of grain growth.

At longer wavelengths, the sensitivity of the JWST will enable extinction curves to be determined with significantly greater precision and across a wider range of environments than was possible in previous work. Both the slope of the infrared continuum extinction and the shapes of the silicate absorption profiles are sensitive to the relative numbers of large grains. By such means, it may be possible to trace the cycle of grain growth, from diffuse to dark clouds, from dense cores to protoplanetary disks.

10.2.2 Ices and Organics in the Interstellar Medium and Protoplanetary Disks

- Can the inventory of interstellar ices in dense clouds be extended, e.g., by detection of O_2, N_2, HDO, and complex organics? What are the limits to complexity in the ices?
- How are ices and organics distributed in protoplanetary disks? Do their distributions vary from one system to another and does this relate to planet formation?
- Do some volatiles survive as solids in diffuse regions of interstellar space, e.g., in very large grains or “cometesimals”?

The importance of infrared space telescopes such as ISO and Spitzer to our current understanding of icy materials in space was noted previously (Section 10.1)—indeed, data from these missions adorn many figures in this book. The JWST promises to take this to the next level. The essential requirements are (i) full coverage of the spectral range containing the relevant solid-state features (Table 5.2), (ii) high sensitivity to enable study of faint sources, and (iii) high spectral and spatial resolutions. Previous missions checked some of these boxes; the JWST and its suite of instruments will check all of them.

Different types of source are used to study different environments in a molecular cloud, as illustrated in Figure 5.2. Background field stars that enable sampling of the quiescent regions are inherently fainter than embedded young stars. The unprecedented sensitivity of the JWST will yield spectra of background stars of a quality previously attainable only for the brightest YSOs. Adequate spectral resolution is also essential for both types of source. Although solid-state features are generally broad they often contain diagnostic structure, and overlapping gas-phase lines could mimic such structure if not fully resolved. Figure 5.13, for example, illustrates the inherent complexity of the CO ice feature and the simultaneous presence of overlapping gas-phase interstellar lines. Spectral lines intrinsic to a source may also contaminate its interstellar spectrum. The resolving powers required to overcome these issues ($\lambda/\Delta\lambda > 10^3$) are baked into the JWST near and mid-infrared spectrometers.

The homonuclear molecules O_2 and N_2 are predicted and potentially abundant constituents of the ices that have so far eluded detection. They produce no spectral features in their pure state; but weak signatures may be induced if, as expected, they are mixed with other species in the ices, such that interactions with near neighbors perturb the symmetry of the molecule. The fundamental O=O vibrational mode of O_2 centered near 6.45 μm, for example, may be a good candidate for detection by the JWST (searches with ISO data yielded modest upper limits). O_2 and N_2 are expected to reside primarily in the CO-rich layer of the ices (Figure 5.14). The presence of these species may induce detectable substructure in the profile of the 4.67 μm solid CO absorption feature, as the result of intermolecular forces.

Deuterium fractionation is recognized as a key signature of low-temperature chemical processes that give insight into the origins of protoplanetary matter in our solar system (Section 9.3.7). Molecular precursors of this material have been detected in the ISM, but so far only in the gas phase. Identification of deuterated water (HDO) and perhaps other deuterated species in the ices would provide important corroboration of the chemical models. The O–D bond in HDO produces a spectral feature at 4.1 μm (equivalent to the strong O–H feature at 3.0 μm). Again, this has eluded detection in previous work, most likely due to insufficient sensitivity, and it thus provides another opportunity for the JWST.

So far, the most complex organic molecule confirmed as a common constituent of the ices (CH_3OH) contains just six atoms. Species of significantly greater molecular weight are known to be present in the gas phase, both in cold clouds and in warmer regions around new-born stars where sublimated mantles may contribute to the inventory (Section 9.3.4). Chemical process in these environments influence the complexity of molecules delivered to protoplanetary disks. In order to refine the models, it is important to test the limits to molecular complexity in the solid phase, and to link this to the gaseous chemistry of hot cores and corinos. The spectral region from 5 to 8 μm is especially important in this regard, as it is populated by the vibrational modes of many candidates for detection (Section 5.3.5). The features observed to date are blends of numerous overlapping absorptions, only a few of which have been identified with confidence. Gains in sensitivity and spectral resolution may enable us to solve this puzzle, not only enhancing the quality of the data but also expanding the range of environments that can be studied. Candidates for detection include acetaldehyde, ethanol, dimethyl ether, and methyl formate, all of which are known to be present in the gas phase in hot cores. The presence of such species in low-temperature ices would inform our models of cold chemistry; or if they are present only in warmer ices, this would imply the need for an energy source to drive production.

The JWST will also enable direct study of ices in protoplanetary disks. Systems in which the disk is aligned approximately edge-on to our line of sight are most suitable, as this configuration alleviates ambiguities concerning the location of the ices (disk versus envelope). Infrared spectroscopy of a spatially resolved disk is not limited to the line of sight to the central star. Spectral imaging may enable entire disks to be mapped in key spectral signatures, seen in absorption against the continuum emission from warm dust. It may thus be possible to map the radial and vertical distributions of the ices. In combination with data on the distributions of gas-phase molecules observed (e.g.) by ALMA, such observations will clarify the interplay between key planet-building materials, and identify the locations of snowlines for solids of different volatilities. Ultimately, the results will inform models for planet formation and planetary-system architecture. In the case of our own planetary system, the orderly separation of rocky planets and gas giants is attributed to formation within and beyond the H_2O snowline of the solar nebula. It will be interesting to determine whether other systems are consistent with this picture, and to construct plausible chemical and dynamical models that can accommodate the wide variety of exoplanetary systems now known to exist.

The present-day distribution of comets in and around our solar system is attributed to the gravitational influence of the giant planets. Oort-Cloud comets arrive randomly from solar distances that far exceed the boundaries of the planetary system. It is hypothesized that they were scattered into highly eccentric orbits by resonances or other disturbances arising from gravitational interplay between these planets. It is unknown whether such disturbances are commonplace or dependent on improbable planetary alignments, but it seems reasonable to suppose that icy planetesimals may be ejected routinely from their parent system. Extrasolar visitors to our own system have now been identified (see Section 7.3.3 for discussion and references). An important long-term goal would be to constrain the space density of debris large enough to retain volatiles in diffuse regions of the interstellar medium.

All things considered, there is clearly no shortage of challenging problems to engage the interested student or more seasoned campaigner.