The history of spin ice

To cite this article: Steven T Bramwell and Mark J Harris 2020 J. Phys.: Condens. Matter 32 374010

View the article online for updates and enhancements.
The history of spin ice

Steven T Bramwell and Mark J Harris

Abstract

This review is a study of how the idea of spin ice has evolved over the years, with a focus on the scientific questions that have come to define the subject. Since our initial discovery of spin ice in 1997, there have been well over five thousand papers that discuss it, and in the face of such detail, it must be difficult for the curious observer to ‘see the wood for the trees’. To help in this task, we go in search of the biggest insight to have emerged from the study of spin ice. On the way, we identify highlights and outstanding puzzles, and celebrate the inspirational role that Roger Cowley played in the early years.

Keywords: spin ice, magnetism, vertex models

1. A ferromagnetic analogue of ice

Pauling’s model of water was introduced in 1935 to explain ice’s residual entropy [1, 2, 3]. It rapidly acquired broader significance as a vivid demonstration of how the properties of a highly complex solid could be captured by a deceptively simple statistical mechanical model. Slater [4], Rys [5], Lieb [6] and others generalised Pauling’s approach to found the field of ‘vertex models’, the extensive study of which, over several decades, provided deep insights into the physics of ferroelectrics, quantum systems and phase transitions [7]. Meanwhile, Anderson in 1956 had proposed an antiferromagnetic analogue of Pauling’s model, based on a pyrochlore (or B-site spinel) lattice of antiferromagnetically interacting Ising spins [8]. By the 1990s, ‘frustrated’ antiferromagnetism was becoming popular and Anderson’s antiferromagnetic model was increasingly seen as its paradigm. But extensive investigations of pyrochlore oxides (e.g. [9]) had not found any realisations of Anderson’s model—there seemed to be no ‘model magnets’ in this class.

In our original discovery [10, 11] of the spin ice model—and of Ho2Ti2O7 as a possible experimental realisation—our contention was that Anderson’s model was unphysical, as its ‘up/down’ Ising symmetry violates the cubic space symmetry and trigonal point symmetry of the pyrochlore lattice, a system of corner-linked tetrahedra (figure 1(a)). Yet the minimal antiferromagnetic model that obeyed these symmetries did not map to Pauling’s model—it would give ‘all in—all out’ order [12] with spins along local trigonal axes (figure 1(b)). But, we discovered, maintaining the spins along the local trigonal axes, and reversing the sign of the nearest neighbour coupling to make it ferromagnetic, restored the mapping to Pauling’s model. Hence the reason that no one had discovered a magnetic analogue of ice was that they had been looking in the wrong place: the magnetic analogue of ice was a frustrated ferromagnet, not a frustrated antiferromagnet. It would be no understatement to say that this observation came as a surprise; indeed, in the months after our initial discovery of spin ice both of us witnessed frequently the shock which this counter-intuitive result would cause in audiences, as we presented our findings.

In our original discovery [10, 11] of the spin ice model—and of Ho2Ti2O7 as a possible experimental realisation—our contention was that Anderson’s model was unphysical, as its ‘up/down’ Ising symmetry violates the cubic space symmetry and trigonal point symmetry of the pyrochlore lattice, a system of corner-linked tetrahedra (figure 1(a)). Yet the minimal antiferromagnetic model that obeyed these symmetries did not map to Pauling’s model—it would give ‘all in—all out’ order [12] with spins along local trigonal axes (figure 1(b)). But, we discovered, maintaining the spins along the local trigonal axes, and reversing the sign of the nearest neighbour coupling to make it ferromagnetic, restored the mapping to Pauling’s model. Hence the reason that no one had discovered a magnetic analogue of ice was that they had been looking in the wrong place: the magnetic analogue of ice was a frustrated ferromagnet, not a frustrated antiferromagnet. It would be no understatement to say that this observation came as a surprise; indeed, in the months after our initial discovery of spin ice both of us witnessed frequently the shock which this counter-intuitive result would cause in audiences, as we presented our findings.

1 Author to whom any correspondence should be addressed.

4 Pauling’s degenerate model developed the ice rule concept introduced by Bernal and Fowler [2].

5 Recent work suggests that Ho2Ti2O7, Dy2Ti2O7 should be more precisely classified as ‘inverting magnets’: where strong ferromagnetic interactions are ultimately overcome by very weak antiferromagnetic ones, in the limit of low temperature [13].
In figure 1(b), with an oxide ion at the centre of the tetrahedron, the spin ice model configurations on a tetrahedron are classified:

- 'Two spins in, two out' describes a water molecule, H$_2$O, so 'in–in' or 'out–out' interactions.

- By the spin ice mapping, four low energy 'in–out' interactions and two high energy compromise is 'two spins in and two out', which consists of neighbour pairs of spins on a tetrahedron. The lowest energy metrically frustrated: it cannot be satisfied between all near water ice. The ferromagnetic interaction between spins is geometrically frustrated: it cannot be satisfied between all near proton configuration in Pauling's model of (cubic) water ice. The ferromagnetic interaction between spins is geometrically frustrated: it cannot be satisfied between all near neighbour pairs of spins on a tetrahedron. The lowest energy compromise is 'two spins in and two out', which consists of four low energy 'in–out' interactions and two high energy 'in–in' or 'out–out' interactions. By the spin ice mapping, 'two spins in, two out' describes a water molecule, H$_2$O, so generates the ice rule of two protons close, two far to each oxygen. Assertion of this rule on each of the $6^2$ vertices and hence the Pauling residual entropy $S = R \log(3/2)$ per mole tetrahedra (other examples are given in the text).

To clearly distinguish it from Anderson's model (and other spin models mapping to Pauling’s [7, 11])—which is important in view of their differing properties—it seemed appropriate to designate this newly discovered ferromagnetic analogue of water ice by its own name: the ‘spin ice’ model [10]. In the spin ice model, configurations on a tetrahedron are classified by the number of spins that point 'in' or 'out', as shown in figure 1(b). With an oxide ion at the centre of the tetrahedron, placing a proton on the end of each spin generates an equivalent proton displacement vectors in water ice. Left to right, respectively: single tetrahedra showing the water ice configurations H$_2$O, H$_3$O$^+$ and H$_4$O$_2$$^+$, with analogous spin ice configurations (red arrows indicate ionic magnetic moments or 'spins', large circle the oxide ion O$^-$, small circles the proton H$^+$). Numbers in square brackets: per tetrahedron, there are six possible 2:2 ('two in two out') states, eight possible 3:1 ('three in one out' or 'three out one in') states and two possible 4:0 ('all in' or 'all out') states. These collectively generate the sixteen ($= 6 + 8 + 2$) vertices and hence the '16-vertex model' description of spin ice. Nearly all of spin ice physics is implicit in this figure. For example, applying the rule that every tetrahedron must be a 2:2 vertex gives a number of equivalent ground states that grows exponentially with system size and hence the Pauling residual entropy $S = R \log(3/2)$ per mole tetrahedra (other examples are given in the text).

![Figure 1](image1.png)

**Figure 1.** (a) A fragment of the cubic pyrochlore lattice of corner linked tetrahedra occupied by the magnetic ions in spin ice. There is an oxide ion at the centre of every tetrahedron (spheres). An analogous structural motif occurs in water ice where the mid-points of the oxide—oxide lines of contact form a pyrochlore lattice (cubic ice [8]) or related hexagonal lattice (hexagonal ice Ih). (b) The spin ice mapping [10, 11], where spins in Ho$_2$Ti$_2$O$_7$, Dy$_2$Ti$_2$O$_7$ become analogous to proton displacement vectors in water ice. Left to right, respectively: single tetrahedra showing the water ice configurations H$_2$O, H$_3$O$^+$ and H$_4$O$_2$$^+$, with analogous spin ice configurations (red arrows indicate ionic magnetic moments or 'spins', large circle the oxide ion O$^-$, small circles the proton H$^+$). Numbers in square brackets: per tetrahedron, there are six possible 2:2 ('two in two out') states, eight possible 3:1 ('three in one out' or 'three out one in') states and two possible 4:0 ('all in' or 'all out') states. These collectively generate the sixteen ($= 6 + 8 + 2$) vertices and hence the '16-vertex model' description of spin ice. Nearly all of spin ice physics is implicit in this figure. For example, applying the rule that every tetrahedron must be a 2:2 vertex gives a number of equivalent ground states that grows exponentially with system size and hence the Pauling residual entropy $S = R \log(3/2)$ per mole tetrahedra (other examples are given in the text).

![Figure 2](image2.png)

**Figure 2.** (left) Some of the neutron scattering data that first suggested that Ho$_2$Ti$_2$O$_7$ approximates the spin ice model [10]. The left hand panel shows the unusual temperature dependence of the magnetic Bragg peaks measured in an applied field of 2 T. The two sets of peaks (full and open circles) respectively suggest the $q = 0$ and $q = X$ magnetic structures, shown as (a) and (b) in the right hand panel. The $q = X$ peaks have an intrinsic width that varies with temperature, indicating incomplete correlations among chains of spins perpendicular to the applied field (large arrow). Reprinted figure with permission from [10], copyright (1997) by the American Physical Society.

correlated state that lacks the broken symmetry expected at low temperature, and hence supports unusual excitations and properties.

Just as Pauling’s model of water ice is ‘locally ferroelectric’, the spin ice model retains a locally ferromagnetic character because, despite the frustration, all except the highest energy spin configurations on a tetrahedral plaquette of the pyrochlore structure carry a magnetic moment, as should become clear by looking at figure 1(b). The local ferromagnetism and ice mapping together hold the key to the physics of spin ice. Indeed, it is little exaggeration to state that all of spin ice’s diverse and interesting properties are implicit in figure 1(b). Just as configurational counting implies the Pauling entropy, other examples include: (i) evaluating the magnetic moment and Zeeman energy of each configuration generates the field-temperature phase diagram, (ii) counting the excess electric charge in each tetrahedron and treating this as magnetic pole density gives spin ice’s magnetic monopoles, (iii) allowing the spins to have transverse components or considering the magnetic analogue of proton tunnelling generates quantum spin ice, and so on—detailed discussion of these properties are the subject of this review. In contrast, the abstract mapping proposed by Anderson [8] for the antiferromagnetic Ising model suppresses many of these properties.

The discovery of spin ice was made in the context of our study of the rare earth (R) titanates R$_2$Ti$_2$O$_7$ [10], a seemingly complex series of dipolar magnets that had previously been studied by several authors [14]. Our proposition was that the contradictory experimental properties of Ho$_2$Ti$_2$O$_7$ could only be explained by the spin ice model (see figure 2). It was clear that the required Ising-like spins, pointing along local
tetrahedral axes, would arise from the crystal field states of Ho$^{3+}$ and that the interaction between these spins was ferromagnetic. It seemed to us that the long-ranged dipolar interactions were not relevant to the problem, being entirely accounted for in the demagnetizing correction. This latter conclusion turned out to be only partly true, but the success of our spin ice model in describing diverse properties was undeniable. The most important development was the analogous Dy$_2$Ti$_2$O$_7$ indeed showed the Pauling entropy of approximately $\log(3/2)$ per mole tetrahedra [15] (the analogous demonstration for Ho$_2$Ti$_2$O$_7$ is in [16] and an up-to-date discussion and confirmation of Ramirez et al’s result is given in [17]).

It is a pleasure to record that our interactions with Roger Cowley played a valuable role in this early development of spin ice. One of us (STB) recalls debating with him, as far back as 1993, whether rare earth ions in the pyrochlore structure are going to be more or less anisotropic than transition metal ions: of course it depends on the quantum states, but essentially the intuition of ‘more’ leads to the idea of classical spin ice and the intuition of ‘less’ leads eventually to the idea of quantum spin ice (of which, more below), so both perspectives can be entertained. Following our development of the spin ice concept, Roger also stressed the value of high resolution neutron scattering at the zone centre—he had pioneered work on ferroelectrics with similar correlations [18]—and indeed such measurements have come to play a crucial role in the field [19–21], pushing the physics of spin ice right to its limits, as described below. MJH recalls that it was Roger who first introduced him to the pyrochlores and their relevance for antiferromagnetic couplings of the spin ice model, thus defined, gives a microscopic interpretation to the phenomenological ferromagnetic coupling of the spin ice model, fixing its value to $J_{\text{eff}}$. The near neighbour model is a truncation of the ‘dipolar spin ice’ (DSI) model [25–27], where further neighbour dipole–dipole and exchange terms ($J_i$) are included. Its Hamiltonian is:

$$\mathcal{H} = \frac{\mu_0 g^2\mu_B^2}{4\pi} \sum_{i<j} \mathbf{S}_i \cdot \mathbf{S}_j - 3 \left( \hat{r}_{ij} \cdot \mathbf{S}_i \right) \left( \hat{r}_{ij} \cdot \mathbf{S}_j \right) - \sum_{i,j} J_i \mathbf{S}_i \cdot \mathbf{S}_j$$

(1)

where $\mu_0$ is the vacuum permeability, $\mu_B$ the Bohr magneton, and $\mathbf{S}$ the spins. The $g$-factor for Ho$^{3+}$, Dy$^{3+}$ takes that value $g \approx 20$ if one sets $S = 1/2$, but it has become practice in the field to set $S = 1$, in which case it is reduced to $g \approx 10$, and we shall continue with the convention subsequently; in either case, the spins are two-state and Ising-like and the magnetic moment is approximately $10 \mu_B$ per ion. In its earliest incarnation, dipolar spin ice included one near neighbour antiferromagnetic exchange term and the full classical dipole–dipole interaction, but to get accurate quantitative agreement with diverse experiments, and particularly with

7 Note that the suggestion in this paper [24], that Ho$_2$Ti$_2$O$_7$ is ‘not spin ice’ proved to be incorrect [16].

8 As written, we have expressed the Hamiltonian in terms of real spin operators: more commonly it is expressed in terms of pseudo-spin operators, with the introduction of appropriate numerical factors [23].
Figure 3. (a) The energy levels of the spin ice 16 vertex model and how they may be tuned by applied magnetic field, chemical substitution and applied strain to create new ground states. (b) Shift and temperature dependence of the excited vertex energies in the monopole model, using parameters appropriate to Dy$_2$Ti$_2$O$_7$. Here the energies were calculated by Debye–Hückel theory for magnetic monopoles, following the method of [69]. The arrows indicate the shift in the energies compared to the near neighbour spin ice model [23].

From one perspective, this question was ultimately answered by the development of the concept of magnetic monopoles in spin ice. Specifically, the monopole model makes clear how dipolar spin ice model approximates near neighbour spin ice in a renormalised theory and quantifies the accuracy of this approximation. But it has recently come to light that there remains a most important case where the near neighbour model actually outperforms the dipolar spin ice model as a description of experiment: why this happens constitutes an unsolved mystery in the physics of spin ice. We explain all this subsequently, but first it is useful to outline some of the successes of the original spin ice model as a guide to experimental properties.

- Q2: why does the near-neighbour spin ice model work as well as it does?
3. Spin ice as a 16 vertex model

Following figure 1, spin ice is a magnetic 16-vertex model (i.e. the sum of six 2:2 vertices, eight 3:1 vertices and two 4:0 vertices), where the Boltzmann weights of the 16 vertices can be tuned by external variables such as temperature, applied field [33], electric field [34, 35], pressure [37] and strain [38] as well as by chemical substitution [39] and quantum effects [40]. This gives rise to an extremely rich thermodynamic phase space, containing disordered, partly ordered and fully ordered phases, separated by many types of phase transition (see figure 3(a)).

At high temperature all 16 vertices are equally weighted and the system is paramagnetic. As the temperature is lowered, higher energy vertices are frozen out: first 4:0 vertices and then 3:1 vertices, leaving the six 2:2 vertices that form the Pauling, or six-vertex state below T ∼ 1 K.

It was quickly realised that a uniform magnetic field applied to the Pauling state can lower the energies of particular vertices, or particular combinations of vertices, leading to several field-induced phase transitions [33]. These may occur by re-weighting states on the 2:2 Pauling manifold, or by stabilising competing 3:1 states that carry a significant [111] magnetic moment. The states formed are easy to infer by a simple consideration of the vertex energies involved (see figure 3(a)), but the nature of the transitions themselves is very subtle and can only be understood by a detailed analysis, and in some cases, by the inclusion of dipole–dipole interactions.

Field induced phases and phase transitions have principally been studied by magnetometry, neutron scattering, analytically and by numerical techniques. The main results may be briefly summarised as follows. An early and notable discovery was that modest fields (∼0.1 T) along [111] stabilise a two dimensional state on the Pauling manifold, with modified residual entropy, so called ‘kagome ice’ [42]. A stronger field (∼1 T) selects a 3:1 ordered state in a transition that is accompanied by a point of high degeneracy, as signalled by a spike in the entropy [43]. At lower temperature the transition evolves into a line of first order transitions with a critical end point [44]. Fields slightly misaligned from [111] give rise to a complex phase diagram, including rare examples of the so-called Kasteleyn transition [45, 46]. Another Kasteleyn transition is observed with a field applied along [100] where an ordered 2:2 state is stabilised [41]. A field applied along [110] gives rise to a partly ordered state containing disordered Ising-like chains [47].

Chemical variation may be used to alter the relative vertex energies in spin ice materials [39, 48]. In particular, if the near neighbour antiferromagnetic exchange is strong enough, the overall interaction becomes antiferromagnetic and the ordered, ‘all in all out’ ground state is stabilised (figure 3(a)). This state is observed in Nd2Zr2O7, for example [49]. Applied strain may similarly alter the relative placing of energy levels (figure 3(a)). In this context, the recent development of epitaxial spin ice thin films [38, 50–52] has added a new dimension to the study of the 16 vertex physics. The defining feature of few-monolayer films of Dy2Ti2O7 grown on nonmagnetic Y2Ti2O7 is homogenous epitaxial strain. This splits the Pauling manifold of spin ice and gives a magnetic realisation of the two dimensional ‘F-model’ with an unusual type of Berezinskii–Kosterlitz–Thouless transition [38].

4. Magnetic monopoles in spin ice

In near neighbour spin ice, the transition from kagome ice to the ordered 3:1 state, as an applied field is increased along [111], remains smooth, no matter how low the temperature becomes (so long as it remains finite) [33]. But in the real material, it becomes a line of first order phase transitions, with a critical end point [44]. This is a clue that there is a correction to the 16 vertex physics, most likely coming from the neglected long range part of the dipolar interaction in equation (1). This, and other dipolar corrections to the physics of the simple 16-vertex model, are almost entirely accounted for by a single beautiful concept: that of emergent magnetic monopoles [55, 56].

The discovery of magnetic monopoles was a very important development that brought spin ice to the attention of a wider community. It had long been appreciated that 3:1 defects in spin ice are analogous to ionic defects in water ice (figure 1(b)) and would play an important role in the magnetic dynamics of Ho2Ti2O7 and Dy2Ti2O7 [54, 76][11], but it was not obvious how to handle the dipole–dipole interaction for these excited states: the natural fear was that any treatment of the dipole–dipole interaction would turn the rather beautiful spin ice physics into something much less attractive: but the opposite proved to be the case.

Although most scientists heard about magnetic monopoles in spin ice through the landmark nature paper of Castelnovo, Moessner and Sondhi (CMS, 2008) [55], the idea was already in the literature, in an important work by Ryzhkin (2005) [56]. The two approaches of CMS and Ryzhkin were quite different, but each of them laid the foundation of a great deal of future work on spin ice, and beyond.

The essential difference was that Ryzhkin adopted the ice mapping and then applied the Jaccard theory of ice defects [57], while CMS started with the DSI and transformed the DSI Hamiltonian into that of a Coulomb gas via the so-called ‘dumbbell model’. Both approaches found that the long range part of the dipole–dipole interaction amounted to a magnetic Coulomb interaction between 3:1 defects[12]. This is perhaps obvious with the remarkable benefit of hindsight (see [58]).

10 The Pauling state in dipolar spin ice is not, in general, completely degenerate as assumed in the monopole model. In dipolar spin ice models, the weak splitting of the Pauling manifold gives rise to antiferromagnetic ordering at low temperature [29], but this is not observed in experiment [17]. The non-degenerate dipolar Pauling manifold generates some exotic physics that has been explored in the analogous case of artificial spin ice [53].

11 The role of such defects in the dynamics of more general ice rule systems is discussed in [54].

12 Note that Ryzhkin and CMS differ over the possible existence of an extra, entropic, Coulomb interaction on top of the magnetic Coulomb interaction.
figure 1(b)): the 3:1 defects are lattice divergences in the magnetization, so constitute magnetic charge as defined in the textbooks, \( V \cdot H = -V \cdot M \). The excited 3:1 and 4:0 vertex states become monopoles of single charge \( \pm Q \) and double charge \( \pm 2Q \) respectively, where \( Q = 2\mu/a \). Here \( \mu \) is the rare earth magnetic moment and \( a \) is the distance between tetrahedron centres, which form a diamond lattice. CMS showed that the [111] phase transition could be treated as a crystallisation of single charge monopoles, while Ryzhkin developed a theory of dynamical response based on neglecting the Coulomb interaction (this was restored in later work [58]). In 2009 a number of works appeared that gave experimental confirmation of these ideas [20, 59–62].

Such developments incidentally closed a circle in ice physics that had begun well before the discovery of spin ice. In 1984 Ryzhkin had constructed pseudo-spin hamiltonians for water ice, equivalent to the dipolar spin ice model [66]. Taking into account CMS’ result and the earlier results that had understood spin ice’s ‘self-screening’, essentially confirmed the equivalence of the pseudo-spin approach to the standard defect-based approach to water ice [57] (this important achievement seems to have gone unheralded in the water ice literature). It was then clear, as Ryzhkin has stressed, that spin ice and water ice have the same properties because they share the same Hamiltonian. In this context, the only significant difference between water ice and spin ice is that, in addition to ‘ionic defects’, water ice has ‘bonding’ defects (two protons or no protons per O–O contact). These are thought to play an important role in water ice’s proton conduction as they can screen the ionic defects and allow a direct current [57].

Although analogues of bonding defects are not possible in the classical models of spin ice, and are not expected in Ho\textsubscript{2}Ti\textsubscript{2}O\textsubscript{7}, Dy\textsubscript{2}Ti\textsubscript{2}O\textsubscript{7}, we remark that they may be possible in some real spin ice materials, where quantum fluctuations can instantaneously replace a dipole by a quadrupole, for example—naively, a quadrupole could behave as a double headed arrow, or no arrow. We return to this idea below, when we discuss neutron scattering, as well as quantum spin ice.

5. Monopole thermodynamics

Monopole thermodynamics may be most elegantly represented in the grand canonical ensemble, with monopole density controlled by a fixed chemical potential, \( \mu = -5.7, -4.35 \) K for Ho\textsubscript{2}Ti\textsubscript{2}O\textsubscript{7}, Dy\textsubscript{2}Ti\textsubscript{2}O\textsubscript{7} respectively [67]. Although this looks like a change of perspective, it remains useful to think in terms of the analogous scheme of chemical equilibria:

\[
2\text{H}_2\text{O} = \text{H}_2\text{O}^+ + \text{OH}^- = \text{H}_4\text{O}^{2+} + \text{O}^{2-},
\]

where the grand canonical ‘vacuum’ is the ice (H\textsubscript{2}O) state with the Pauling entropy. However, a number of differences with the usual chemical algebra must be noted. First, the various quasi-chemical species do not have equal weighting at infinite temperature. At the level of the Pauling approximation, the weighting is 6:4:4:1:1 (reading right to left in equation (2)). Second, the constraint of time-reversal symmetry guarantees that +1 charges have the same concentration as −1 charges, and that +2 charges have the same concentration as −2 charges at equilibrium, which is formally guaranteed by the simplified equation (2), but not necessarily in the most general chemical analogy. This means that some care must be exercised in applying the quasi-chemical analogy at a quantitative level.

It is nevertheless possible to use textbook methods of chemical thermodynamics to calculate the equilibrium thermodynamic properties of spin ice. In general Coulombically-interacting charges, such as ions in solution, or magnetic monopoles in spin ice, need to correlate strongly in order to make a finite internal energy and regularise thermodynamics (randomly placed charges have infinite energy in the thermodynamic limit). The controlled theory of this effect, the Debye–Hückel theory, shows that the bare Coulomb potential of a single ion is exponentially screened at equilibrium by the formation of a surrounding ‘ionic cloud’ of oppositely charged ions. The effective potential then takes the form \( (1/r) e^{-r/\ell_D} \) where \( \ell_D \) is the Debye length. In practical terms, this thermodynamic effect of the Coulomb energy may be absorbed into an activity coefficient \( \gamma_i \) (where \( i \) denotes an ionic species) and this is what Debye–Hückel theory calculates.

A Debye–Hückel theory of \( \gamma_i \) in spin ice may be developed in exactly the same way as for ionic solutions [68] (here \( i = 1, 2 \) indicate single and double charge monopoles respectively). The result is [69]:

\[
\log(\gamma_1) = -\frac{l_T}{\ell_D + a}, \quad \log(\gamma_2) = -\frac{4l_T}{\ell_D + a},
\]

where \( l_T \) is the Bjerrum length for single charge monopoles and \( \ell_D \) is the Debye length, as usually defined [68]. In the magnetic context, these parameters are given by\(^{14}\):

\[
\ell_D = \sqrt{\frac{\mu_i Q^2(n_1 + 4n_2)}{kT v_0}}, \quad l_T = \frac{\mu_i Q^2}{8\pi kT},
\]

where \( n_1, n_2 \) are the number density of single and double charge monopoles respectively, per site of the diamond lattice that they inhabit and \( v_0 \) is the volume per site. The difference with ionic solutions (apart from a change from electrical to magnetic variables) arises from the spin ice correlations already present before the Coulomb interaction are switched on. It may be accurately (though not exactly) accounted for by replacing the ideal solution reference state by the Pauling state, with effective chemical potential \( \mu_i = \mu - kT \log \gamma_i \). The Pauling approximation is equivalent to a single-tetrahedron or single-vertex approximation but with exactly half the number of tetrahedra that are actually in the pyrochlore lattice (this is because the lattice has two spins per tetrahedron, rather than four). Hence we find for the concentration of monopoles per

\(^{13}\) The \( \mu \)SR technique used for the experiment described in [62] was questioned by [66]. See [67] for a summary of the issues. The experiment was later repeated unambiguously by [68] using magnetometry.

\(^{14}\) Note that [58] defines the Debye length slightly differently: the difference is connected to the fact that the potential is not purely of the Debye form at long range, but this has negligible effect on the energy.
site of their lattice:

\[ n_1 = \frac{8\tilde{\mu}_1}{6 + 8\tilde{\mu}_1 + 2\tilde{\mu}_2} \]
\[ n_2 = \frac{2\tilde{\mu}_2}{6 + 8\tilde{\mu}_1 + 2\tilde{\mu}_2} \]

Because the \( \gamma \)'s and the \( n \)'s are mutually dependent, to solve these equations numerically requires an iteration, but that is easily done. Having found the densities, \( n_1(T), n_2(T) \), the specific heat is easily calculated by standard thermodynamic methods ([69], equations (7)–(10)). As is well known, the Debye–Hückel approximation needs to be supplemented by a ‘Bjerrum correction’ that accounts for short-ranged dipole-like correlations between charges, a more complex calculation that is achieved in [69]. The Debye–Hückel–Bjerrum theory, thus formulated, gives an entirely convincing analytic description of the specific heat versus temperature of the spin ice materials (figure 4).

Debye–Hückel approaches to spin ice were first described in [60] and later in [39, 70], but a complete analysis is only given in [69], which supersedes the previous approaches. Although Debye–Hückel–Bjerrum theory is often seen as a ‘low density’ approximation, it really applies for the small parameter \( n/T \); the temperature factor ensures that it gives an accurate approximation to spin ice at even quite high temperatures.

6. Answer to question 2

We are now in a position to give our first answer to question 2, as defined above: why does the near neighbour spin ice model work as well as it does? We have shown how, in zero applied field, the monopole fluid of spin ice acts, to an excellent approximation, as near neighbour spin ice with an effective near neighbour energy scale of \( |\mu| \) rather than \( J_{\text{eff}} \).

The result is plotted in figure 3(b) where it can be seen that the primary effect of the Coulomb energy is simply to renormalise \( J_{\text{eff}} \) to a significantly larger value. In fact, for the purpose of using near neighbour spin ice to describe the specific heat, it would already be a far better approximation to replace \( 2J_{\text{eff}} \) with \( |\mu| \) where the two are related by [75]:

\[ |\mu| = 2J_{\text{eff}} + \sqrt{2}C(4\sqrt{6} - 3)/6\pi = 2J_{\text{eff}} + 0.51C. \]

Here \( C \approx 4 \) K (for both Ho₂Ti₂O₇ and Dy₂Ti₂O₇) is the Curie constant, so the near neighbour approximation is seen to underestimate the 16 vertex energy parameter by about 2 K, a large correction to \( 2J_{\text{eff}} \). Although the correction is large, it may be seen from figure 3(b) that dipolar spin ice quite well approximates a near neighbour spin ice model with \( 2J_{\text{eff}} \) replaced by \( |\mu| \). However, this argument only applies in zero field: in a strong enough applied field the Debye screening will be suppressed and \( J_{\text{eff}} \) will become the more relevant parameter. So, strictly, dipolar spin ice approximates the 16 vertex model only if near neighbour coupling is allowed to be temperature dependent (as in Debye–Hückel theory) and field dependent.

The field dependence of the relevant energy scale—the chemical potential—may be quantified by absorbing the Zeeman energy into a ‘magnetochemical potential’, somewhat analogous to the well-known electrochemical potential of electrochemistry and semiconductor physics. Thus a [111] field is seen to reduce the chemical potential for single charge monopoles, such that they eventually proliferate in large numbers in a one tesla field, causing the ‘monopole crystallisation’ described by CMS [55] (see above).

Similarly, certain chemical replacements (e.g. Ti for Ge [39]) can be seen to reduce the absolute value of the chemical potential (for example by increasing the strength of antiferromagnetic superexchange) and enhance monopole creation at low temperature, leading to stronger correlations and eventually crystallisation into a double charge monopole crystal: an ‘all in, all out’ (4:0) antiferromagnetic state.

As realised by the authors of [40] in 2014, the replacement of fixed couplings by monopole chemical potentials creates a yet richer set of possibilities than offered by the dipolar spin ice model. Thus, in dipolar spin ice, the double charge monopole chemical potential, \( \mu_2 \), is always four times the single charge one, \( \mu_1 \). But if the two are unlinkd such that, for example, \( |\mu_2| \gg |\mu_1| \), there arises physics of great subtlety and interest. For example, the underlying magnetic moments can ‘fragment’ into an disordered, spin ice-like part and a crystalline, charge-ordered part. This is unlikely to happen in the classical spin ice materials that closely approximate the dipolar spin ice model, but it has been observed by neutron scattering on spin ice systems where quantum fluctuations are more important, such as Nd₂Zr₂O₇ [71]. For the latest results concerning this interesting development, we refer to [72].

7. Susceptibility and neutron scattering

The Pauling ground state manifold contains no lattice divergence (i.e. at each point there are as many arrows pointing into
a vertex as are pointing out—see figure 1(b) and is the vacuum for magnetic monopoles. For this reason it is referred to as the ‘Coulomb phase’ [73]. The truncation of the solenoidal lattice field lines at the system boundaries gives harmonic fields (fields devoid of divergence and curl), which carry magnetization. The thermal development of these harmonic modes, which become topological objects for periodic boundaries, are manifest as a ‘Curie Law crossover’ in the bulk susceptibility and are called ‘topological sector fluctuations’ [74]. To see this Curie law crossover in experiment requires very accurate demagnetizing corrections [75].

Turning to neutron scattering, reference [58] has provided a calculation of the dynamic scattering function $S(Q, \omega)$ (and associated generalised susceptibility) of the monopole model. However, it is not possible to test the dynamical part (i.e. the spectral shape function) by neutron scattering because the relaxation in the canonical spin ice materials is too slow to clearly resolve the line shape at low temperatures, although neutron spin echo can see some relaxation [76]. Therefore, attention has focussed on the static structure factor for spin or magnetization fluctuations, $S^{\alpha\beta}(Q)$, which as a result of developments in neutron scattering instrumentation is typically imaged over large slices of reciprocal space, a characteristic signature of the field [20, 21, 77].

The static structure factor tensor $S^{\alpha\beta}(Q)$ is a quantity of great interest as it is the Fourier transform of the two-spin correlation function. It is available to an inelastic (triple axis) neutron scattering experiment by integrating over energy transfers. However, in the case of spin ice, energy transfers are sufficiently small that the integral is performed to a good approximation in an ordinary double-axis experiment (the ‘static approximation’). The quantity measured, the differential cross section, is proportional to $S^{\alpha\beta}(Q)$ projected transverse to the scattering vector $Q$:

$$\frac{d\sigma}{d\Omega} = A(Q) \left( \delta_{\alpha\beta} - \hat{Q}_\alpha \hat{Q}_\beta \right) S^{\alpha\beta}(Q).$$

Here, $\sigma$ is cross section, $\Omega$ is solid angle, hats represent unit vectors and $A(Q)$ contains constants and the ionic form factor. An unpolarised neutron scattering experiment sums over the components $\alpha, \beta = x, y, z$ and hence information on correlations may be lost in the summation. However, in most cases, the tensorial character of the structure factor is rather trivial. For example at an ordinary critical point in a cubic system, neglecting dipolar effects, the system has rotational invariance and equal eigenvalues of the structure factor. In that case the full tensor may easily be reconstructed from unpolarised measurements, and this allows unpolarised neutron scattering to be a very powerful probe of magnetic correlations.

Spin ice is an exception to this particular rule. Although the dipolar spin ice model with up to third nearest neighbour exchange couplings provides an adequate description of the unpolarised neutron scattering in zero applied field [27] (figure 5), the unpolarised neutron scattering does not contain the full information on spin ice’s subtle magnetic correlations. The reason for this is that the ice rules and dipolar interactions (or monopoles) conspire to break the rotational invariance of the magnetic correlations. The structure factor tensor is characterised by two different eigenvalues: one longitudinal and one transverse to the wavevector $q = Q - G$ in each Brillouin zone defined by reciprocal lattice vector $G$. It requires neutron polarisation analysis to separate these components (figure 6), which was first achieved in 2009 [20]. The experiment from [20] yields a spin flip (SF) and non spin flip (NSF) cross section. The NSF cross section measures correlations out of the scattering plane defined by $Q$ and hence, from equation (3), gives the pure transverse eigenvalue. The SF cross section gives a mixture of transverse and longitudinal eigenvalues within the scattering plane.

Figure 5. Diffuse unpolarised neutron scattering from Dy$_2$Ti$_2$O$_7$ fitted to the generalised dipolar spin ice model with first, second and third nearest neighbour exchange constants: the parameter values were initially inferred from thermodynamic measurements and it can be seen that they give a consistent description of the neutron scattering data. The Dy in the sample was isotopically enriched to reduce absorption [28]. Reprinted figure with permission from [27], copyright (2008) by the American Physical Society.

A most striking feature of the scattering patterns of figure 6 are the ‘pinch points’ seen in the SF channel, and the overall lack of periodicity of that pattern (it has point symmetry only). The two eigenvalues of $S^{\alpha\beta}(Q)$ are periodic with the reciprocal lattice and so do not contain pinch points, as seen in the NSF cross section of figure 6. The pinch points arise from certain projections of the tensor, one such being delivered by the transverse projection operator of equation (3). Individual tensor components like $S^{xxy}$ also contain pinch points. Ultimately the pinch points simply reflect the difference between longitudinal and transverse eigenvalues. Physically, this arises because there is an energy cost associated with divergences of the magnetization $M$, arising from both the ice rules (see figure 1(b)) and more subtly, the long range dipole–dipole interaction. This tends to suppress the longitudinal eigenvalue relative to the transverse. The pinch points thus serve to emphasise the breaking of rotational invariance by the local ice rules and long range dipolar interactions.

Pinch points are to some extent a diagnostic of the Coulomb phase. The ice rule ‘two spins in, two out’ acts to suppress lattice divergence of the magnetization (see figure 1(b)), ensuring $\nabla \cdot M = 0$ and zero magnetic pole density. The magnetization flux is then pure solenoidal and only the transverse eigenvalue is finite, leading to sharp pinch points. The ice rule
Figure 6. The spin correlations in spin ice are fully revealed only by polarised neutron scattering. The upper figure shows the polarisation analysed scattering patterns—spin flip (SF), non spin flip (NSF) and total—of Ho$_2$Ti$_2$O$_7$. Note that the spin flip scattering shows narrow necks or ‘pinch points’ at 002 (top) and 111, 222 (diagonal) and symmetry-related positions. Lower figures show scans across the pinch point at 002, comparing different components (left) and illustrating the temperature dependence of the SF profile (right).

thus implies a local gauge symmetry that creates a conserved quantity—the solenoidal magnetization flux [73]—and the pinch points of near neighbour spin ice [20] reflect this. Excitation of monopoles gives \( \nabla \cdot \mathbf{M} \neq 0 \), hence an irrotational component of the magnetization: at small \( q \) the equality of transverse and longitudinal eigenvalues is restored and the pinch points broaden.

However, this simplified argument [73] does not apply to dipolar spin ice. In a geometric picture, the inverse structure factor \( S^{\alpha\beta}(Q) \) may be represented as an ellipsoid with principal axes given by the square roots of the transverse and longitudinal eigenvalues [78]. With the suppression of the longitudinal component in the Coulomb phase (absence of monopoles), the tensor becomes a disk, but in fact sharp pinch points occur for any degree of eccentricity, and as the dipole–dipole interaction gives an eccentric tensor as \( q \to 0 \), the pinch points may exist at all temperatures (this has been called the ‘harmonic phase’, because the thermally averaged fields can be described as harmonic [78]). Thus in the dipolar spin ice model, the excitation of magnetic monopoles does not broaden the pinch points [79]. The physical reason is that the classical monopoles are incapable of fully screening each other at long distances, which leads to a suppression of charge fluctuations as \( q \to 0 \). In general the simulations [79] of the structure factor \( S^{\alpha\beta}(Q) \) of the dipolar spin ice model support the predictions of [58].

Despite this, it is found that in the real materials, Ho$_2$Ti$_2$O$_7$, Dy$_2$Ti$_2$O$_7$, the pinch points are broadened, and quantitatively so, to the precise degree expected if full screening is restored [79]. Essentially the data matches much more closely to the near neighbour model (where defects are free) than the dipolar model, where the monopoles are confined at long distance: this is the further mysterious aspect of ‘question 2’, alluded to above. At present, the origin of this serious discrepancy between theory and experiment is not understood. One possibility is that quantum fluctuations create new channels for monopole screening, in a way somewhat analogous to the way bonding defects in water ice are able to screen the ionic defects and allow the passage of a direct current.

The discovery of spin ice was based in large part on neutron scattering observations [10], so it is ironic that the neutron scattering of Ho$_2$Ti$_2$O$_7$ and Dy$_2$Ti$_2$O$_7$ remains not yet fully understood, after so many years. Roger’s view, that the most important probe of spin ice was high resolution neutron scattering at the centre zone, was indeed prophetic.

8. Monopole currents

The process of conduction in spin ice or water ice materials proceeds by the monopoles or ionic defects migrating through the system in response to an applied magnetic or electric field respectively. The passage of defects causes the growth of polarisation (water ice) or magnetization (spin ice) through the stretching out of polarisation tubes or ‘Dirac strings’ [55] that connect the charges. In water ice, bonding defects associate with, and thereby screen, the ionic defects, such that a non-polarising current can be passed. In classical spin ice there is no analogue of bonding defects, so only a magnetising current is possible, and this cannot persist in a closed loop or toroidal geometry. The reason is in part that the magnetization blocks channels of conduction, but more importantly, it reduces entropy, leading eventually to thermodynamic equilibrium. Defining the current density as the rate of change of magnetization with time, \( \mathbf{J} = \partial \mathbf{M} / \partial t \), the latter effect is summarised in the following conduction equation [56]:

\[
\frac{\partial \mathbf{M}}{\partial t} = \kappa \left( \mathbf{H} - \frac{\mathbf{M}}{\chi} \right). \tag{4}
\]

Here \( \mathbf{H} \) is the applied field, \( \chi \) the susceptibility and \( \kappa \) is a monopole conductivity, which is linear with monopole density. The right hand term is an entropic reaction field analogous to the Jaccard (polarisation) field in water ice [57]. This term originates in the entropy cost of the passage of monopoles which magnetizes the system. The free energy density cost of magnetization is \( F = \mu_0 \mathbf{M} \cdot \mathbf{M} / 2 \chi \) and by writing \( \partial \mathbf{M} / \partial t = -(\kappa / \mu_0) \partial F / \partial \mathbf{M} \) (i.e. by treating the magnetization as a thermodynamic displacement from equilibrium) one recovers the right hand term in the brackets of equation (4) (a microscopic derivation is given in [56]).

The study of magnetic relaxation in spin ice has a long history: early studies evidenced collective effects [80] and quantum tunnelling [76, 81]. In general, the monopole...
concept rationalised such effects [59] and more recently, the key equation (4) has been tested and confirmed [82].

The magnetization and field in equation (4) may be treated as local variables and generalised to include diffusion terms and a ‘screening field’ that arises from other monopoles in the system. Then combination with the Poisson and continuity equations gives a complete set of drift–diffusion equations for monopole motion and response [88]. These are analogous to the Nernst–Planck–Poisson equations of electrochemistry or the Van Roosbroeck equations that describe semiconductor response on mesoscopic scales [83], the only difference (apart from a change from electric to magnetic variables) being the inclusion of the entropic reaction field. The full set of equations may be written [58]:

\[
j_{\pm} = -D_{\pm} \nabla n_{\pm} + \frac{Q_{\pm} D_{\pm} n_0}{kT} \left( H + h - \frac{M}{\chi} \right) \tag{5}\]

\[
J = \sum_{\pm} Q_{\pm} j_{\pm} \tag{6}
\]

\[
\frac{\partial n_{\pm}}{\partial t} = -\nabla \cdot j_{\pm} + \ldots \tag{7}
\]

\[
\nabla \cdot h = \sum_{\pm} Q_{\pm} \delta n_{\pm} \tag{8}
\]

where \( j_{\pm} \) is the flux density of each species (\( \pm \)) of charge \( Q_{\pm} \), \( D_{\pm} \) are the diffusion constants \( (D_{\pm} = D) \), \( \delta n_{\pm} \) are the deviations from the equilibrium concentration \( n_0 \) and \( h \) is the screening field. The conductivity \( \kappa = 2DQ_{\pm} n_0/kT = 2u n_0 \) where \( u \) is the monopole mobility. Equation (6) defines the magnetic current density, and equation (8) is the magnetic analogue of the Poisson equation. Equation (5) gives the magnetic equivalent of the Nernst–Planck equation in the limit \( \chi \rightarrow \infty \). Equation (7) is the continuity equation, where carrier generation and recombination terms, denoted by \( \ldots \), are omitted (inclusion of these terms gives a nonlinear response, discussed below).

From a magnetic perspective the response functions of most interest are the dynamic screening function tensors \( S^{\alpha\beta}(Q, \omega) \) and \( S^{\alpha\beta}(Q, t) \). The following explicit predictions for these functions may be derived from equations (5)–(8) [58]:

\[
S^{\alpha\beta}(Q, \omega) = 4Q^2 D n_0 \left[ \frac{1}{\tau^{-2} + \omega^2} \left( \delta_{\alpha\beta} - \frac{q_\alpha q_\beta}{q^2} \right) \right. \\
+ \frac{1}{(\tau')^{-2} + Dq^2} \frac{q_\alpha q_\beta}{q^2} \right] , \tag{9}
\]

\[
S^{\alpha\beta}(Q, t) = \left( \frac{\chi T}{3C} \right) \left[ \left( \frac{q_\alpha q_\beta}{q^2} \right) e^{-1/\tau} \right. \\
+ \left. \frac{q_\alpha q_\beta}{q^2} \frac{e^{-1/\tau^' + Dq^2}}{\tau' + \tau} \right] , \tag{10}
\]

where \( \tau = \chi/\kappa \) and \( \tau' = \tau/(1 + \chi) \) are transverse and longitudinal and relaxation times respectively and \( C \) is the Curie constant. The static structure factor \( S^{\alpha\beta}(q) \) may be derived from equation (10) by setting \( t = 0 \). As discussed above, the static scattering function simulated for the dipolar spin ice model is consistent with this prediction, though certain discrepancies with experiment remain to be resolved.

One might ask, is it possible to eliminate the entropic term in equation (5), so that spin ice undergoes normal, Drude-like conduction of magnetic monopoles? We stress that dipolar spin ice cannot support a direct current (as shown by equations (3) and (4)), but a slight modification to the model that makes it fully equivalent to Bjerrum’s model of proton conduction in water ice, might enable a direct current. This would require the analogue of bonding defects in spin ice: a ‘double headed arrow’ or ‘no arrow’ on a pyrochlore lattice site. One might speculate that an ion with a quadrupolar ground state but low lying magnetic (dipolar) excited states could generate the required analogue of bonding defects to allow a persistent current (of course this could still not pass through the boundaries of the material into non-spin ice materials). A number of quantum spin ice candidates indeed have the property of nearby nonmagnetic states [84] (see below).

At temperatures below \( T \approx 0.7 \) K, the number of monopoles diminishes exponentially and spin ice enters a so-called ‘frozen’ state of very slow relaxation [81]. It became clear that any measurements in this regime are only meaningful if states of the system can be reproducibly prepared. Paulsen et al [85] described a magneto-thermal quench method that can be used to accurately and reproducibly prepare non-equilibrium monopole populations for subsequent investigation. This requires fine control of temperature, magnetic field and the timing of measurements. The development of this technique by Paulsen et al [85] opened up considerable possibilities for the investigation of magnetic monopole currents as well as monopole dynamics at low temperatures, as described in the next section.

9. Field assisted charge generation

The generation of Coulomb charges \((+/−)\) from a real or effective vacuum \((0)\) is of relevance to many processes, from the creation of electron–positron pairs from the Dirac sea to the creation of electron–hole pairs in a semiconductor and the auto-ionisation of water. Once generated the free charges \((+/−)\) interact according to Coulomb’s law, \( V = Q^2/4\pi\epsilon_0\epsilon_r \), but this is nearly always conditioned by the nature of the vacuum. If the latter behaves as a classical polarisable fluid then one can simply substitute \( \epsilon_0 \rightarrow \epsilon_0\epsilon_r \), where the relative permittivity \( \epsilon_r > 1 \). More generally, interaction of charges with the vacuum can lead to complex static and dynamic effects, ranging from the effects of ‘vacuum polarisation’ in QED (Lamb shift and g-factor anomaly) to the Jaccard field that affects the motion of ionic defects in water ice.

In most cases there is an initial dipole excitation, which subsequently ‘fractionalises’ into two free charges or monopoles:

\[
(0) = (++−) = (+)+ (−). \tag{11}
\]

To unbind, the charges have to negotiate the Coulomb energy barrier formed by their mutual attraction. This may occur
situation depicted in figure 7(a).

These processes may be assisted by the barrier, or else by quantum mechanically tunnelling through the barrier. These processes may be assisted by the barrier, or else by quantum mechanically tunnelling through it: both give characteristic signatures in the current density. (b) The classical exponential root field form for (upper) thermionic emission from a tungsten filament at \( T = 1918 \) K whereby electrons escape their image charges (data from [94]), and (lower) pair unbinding in spin ice at \( T = 65 \) mK, whereby monopoles and antimopoles escape their mutual attraction (data from [92]—this is the initial, or transient current). The exponential root field form is a signature of the Coulomb law and rules out other forms of interaction [92].

Figure 7. (a) An initially excited pair of charges (e.g. a monopole–antimonopole \((+/-)\) pair in spin ice) sees a Coulomb barrier formed by sum of the applied potential and Coulomb potential. Charges may escape thermally over the barrier or else tunnel through it: both give characteristic signatures in the current density. (b) The classical exponential root field form for (upper) thermionic emission from a tungsten filament at \( T = 1918 \) K whereby electrons escape their image charges (data from [94]), and (lower) pair unbinding in spin ice at \( T = 65 \) mK, whereby monopoles and antimopoles escape their mutual attraction (data from [92]—this is the initial, or transient current). The exponential root field form is a signature of the Coulomb law and rules out other forms of interaction [92].

either by acquiring sufficient thermal energy to classically surmount the barrier, or else by quantum mechanically tunnelling through the barrier. These processes may be assisted by the application of an electric field, in which case one obtains the situation depicted in figure 7(a).

The ensuing current is determined by the instantaneous density of free charges and by their rate of passage across or through the Coulomb barrier. The field-dependence of the current is generally nonlinear, or non-Ohmic, and historically, the first observed exceptions to Ohm’s law were identified in such processes [86]. Departures from Ohm’s law depend in detail on the dynamics involved, the nature of the vacuum and so on, but there are two universal functional forms, applicable at high field, which do not depend on the details of the system. These forms, which may be derived from simplified models, are characteristic of classical emission and quantum tunnelling respectively. They are as follows [87–90]:

\[
I \sim \exp(\beta_1 Q^{1/2} \sqrt{E/T}) \quad \text{(classical)}
\]

\[
I \sim \exp(-\beta_2 \sqrt{mW/QE}) \quad \text{(quantum)}
\]

Here \( E \) is the electric field, \( \beta_1, \beta_2 \) are known constants that depend on the vacuum and the dynamical process, \( W \) is a measure of the barrier height, \( m \) is the particle’s mass and \( Q \) its charge. In deriving the quantum expression, the detailed shape of the Coulomb barrier is approximated by a triangular function [88], but in the classical case the full shape is retained.

Examples of the classical field-assisted unbinding include the Wien effect in electrochemistry [86, 89], the Poole–Frenkel conduction in semiconductors [90] and Schottky field emission from a metal surface [87], whereby an electron escapes its image charge. Examples of the quantum process include Fowler–Nordheim tunnelling [88] and the Schwinger mechanism [91].

The classical process in particular is a sensitive diagnostic of the force law between escaping particles or quasiparticles [92]. Thus, in equation (12), the characteristic scaling with charge, field, and temperature, depends only on the shape of the Coulomb barrier itself—the ubiquity of the Coulomb interaction is the source of the universality. From these considerations it is clear that measurements of departures from Ohm’s law may potentially be used to identify the process (quantum or classical), measure the charge, test the Coulomb law, and determine the rate constants of unbinding and recombination.

The controlled magnetothermal quench method developed by Paulsen et al [85] allowed direct observation of the field assisted unbinding of frozen-in monopole–antimonopole \((+/-)\) pairs at very low temperature (\( \sim 65 \) mK) [92]. Magnetic monopoles are far from thermodynamic equilibrium at this temperature [93], but the theory is a mechanical-kinetic one that, in common with other theories in this class, is well adapted to describing this situation (see also below). The experiments revealed the magnetic version of the universal classical form for magnetic monopoles in spin ice and gave a very direct confirmation of the predicted Coulomb interaction between magnetic monopole defects and a measurement of their charge and current. To emphasise the universality of the phenomenon, we may quantitatively compare these measurements to those on an electrical system that is at first sight very different. To do this we first express the magnetic field and monopole current in electrical units. In figure 7(b) we compare historic data for Schottky emission from a tungsten filament at \( T = 1918 \) K [94] with that measured for magnetic monopoles in spin ice at \( T = 65 \) mK [92]. One can see that, apart from a temperature rescaling, the phenomena are indeed analogous, the antimonopole taking the role of the ‘image charge’ of the monopole, and vice versa. The reason for the vast difference in energy scales between the magnetic

\[\text{Here the monopole interaction } V = \mu Q^2/4\pi r \text{ was replaced by } V = (Q^2/4\pi r)cQ \text{ where } c \text{ is the speed of light. With } Q \text{ measured in coulombs, the current density } (\partial M/\partial t) \text{ is measured in amps/m}^2: \text{ it was converted to current in amps by accounting for the sample dimensions.}\]
Figure 8. Neutron scattering from a candidate quantum spin ice, Pr$_2$Hf$_2$O$_7$ (left), compared to theory for quantum (QSI) and classical (CSI) spin ice (right). The experimental data at 50 mK, from which a background has been subtracted, are integrated over a narrow band of energy transfer near to zero. Reprinted by permission from Macmillan Publishers Ltd: [Nature Physics] [106], copyright (2018).

and electric experiment is the difference in charge. In electrical units the monopole charge is only $e/112$, much less than the electron charge. But it can be seen that the characteristic ‘exponential root field’ conduction is observed in both cases.

In an intermediate temperature range ($\sim$0.5 K), the field-assisted generation and recombination of magnetic monopoles becomes a much more subtle problem, which may be described very accurately by an adaption of Onsager’s theory of the Wien effect for field dissociation in weak electrolytes, a remarkable example of a nearly exact theory of nonlinear and non-equilibrium relaxation in a many body system [95, 96]. Originally developed for liquid electrolyte solutions [89], the theory has been adapted for spin ice to account for the lattice [95] and the Jaccard reaction field [96]. Detailed numerical simulations have shown that at short times, the Wien effect—a linear increase in charge density giving a non-Ohmic current—is just the same as that expected for a symmetric lattice electrolyte [96]. However, at long times, the current is suppressed by the reaction field. This gives a remarkable non-monotonic approach to equilibrium, whereby, after a field is applied, the density of monopoles first increases and eventually decreases again at long times. In addition, the dependence of density on the modulus of field at short times, gives a ‘frequency doubled’ response of density to an applied ac-field. From the ice mapping (figure 1(b)), such processes could be relevant to water ice as well.

10. New horizons

Many of the themes described above continue to motivate spin ice research to this day. A recent development in monopole kinetics has been the experimental study of monopole ‘noise’, somewhat equivalent to ‘shot noise’ in electricity [97]. It would be interesting to achieve this for the pair unbinding processes revealed in figure 7(b). Monopole dynamics, as opposed to kinetics, is a topical subject of research, with experiment revealing strong evidence for tunnelling processes [82], including those assisted by nuclear spins [98], and theoretical approaches based on the crystal field Hamiltonian [99]. There are now many spin ices known [100], the chemical phase space of $\text{A}_2\text{B}_2\text{X}_7$ being extremely large, including those based on spinels rather than pyrochlores [101].

This ability to change the parameters of the problem by chemical variation plays a particularly important role in the study of ‘quantum’ spin ice [102]. The classical spin ice state contains six membered hexagonal spin loops, each with a particular sense of rotation [11]. If these can be made to reverse their rotation by quantum tunnelling, there arises the possibility of a stable, three dimensional spin liquid state [103, 104], and great effort has been expanded in the experimental search for such a state. Theoretically, the recipe for quantum spin ice involves non-Ising terms for individual spins as well as the tuning of weak, higher order, exchange terms in the Hamiltonian [102]. As such subtle effects are hard to control, the only rational strategy for the experimentalist is to interrogate as many quantum spin ice candidates as possible. To this end, there have been some very beautiful neutron experiments on candidate systems: the most promising candidates include Pr$_2$Zr$_2$O$_7$ [105], Pr$_2$Hf$_2$O$_7$ [106], Nd$_2$Zr$_2$O$_7$ [71] and Ce$_2$Zn$_2$O$_7$ [107]. As a complement to the ‘emergent electrostatics’ of classical spin ice, quantum spin ice [102] offers a fully dynamic emergent electromagnetism, replete with artificial photons as well as monopoles (spinons) and visons (sources of fictitious electric field named after ‘Ising vortices’ [108]). The main signatures of quantum spin ice are found in neutron scattering [109]. As an example, we shown in figure 8 inelastic scattering data from [106], which is compared to theoretical predictions of the quantum and classical spin ice models. A careful analysis of the spectrum gives evidence for photon excitations in these materials [106].

The name ‘quantum spin ice’ was first coined in [110] and stressed in an elegant neutron scattering study of Yb$_2$Ti$_2$O$_7$ [111], which brought the idea to wider attention. Once again,
following figure 1(b), there is a strong analogy between quantum spin ice and water ice, where a proton liquid state has been predicted [112]. In analogy with spin ice, the key component in water ice is concerted tunnelling events involving hexagonal loops of proton moves. In fact, there is experimental evidence from incoherent quasielastic neutron scattering that such events do occur [113].

So far, we have not mentioned artificial spin ice. This is a major branch of spin ice physics that was started in a foundational paper by Wang, Schiffer and colleagues, published in 2006 [114]. In artificial spin ice, atomic magnetic moments are replaced by mesoscopic ferromagnetic islands that behave like tiny bar magnets. Because the islands are placed in a plane, the analogy is with two dimensional ice, and direct space probes such as magnetic force microscopy (MFM) can be used to image the magnetic correlations. Square spin ice arrays have similar square plaquette configurations to the ‘in–out’ configurations of figure 1(b), and ‘pole exclusion’ makes the 2:2 set lowest in energy. In the simplest arrays, the Pauling manifold is not degenerate, because the six ‘two in two out’ states in a square geometry are not all symmetrically equivalent (this may be seen by considering the interaction of four bar magnets placed along the arms of a cross). The degeneracy can, however, be restored in two ingenious ways. First ([115], following the idea of [116]) by offsetting the islands in the third dimension to a precisely controlled degree, and second ([117]) by retaining two dimensional geometry, while introducing magnetic disks at the vertex points between islands. These disks act as interaction modifiers such that the degeneracy of the Pauling manifold can be fine tuned [117]. More generally, artificial spin ice enables the engineering of a vast selection of lattice types and geometries, where frustration effects and degeneracies may be finely controlled [120]. The physics of artificial spin ice is now a very large and thriving subject: we cannot do justice to it here, but instead refer to recent reviews [118, 119].

11. Conclusions

Throughout this review, we have referred back to figure 1, because, as we have illustrated, nearly all of the interesting aspects of spin ice physics—the residual entropy, Coulomb phase, topological sector fluctuations, monopoles, fractionalisation, fragmentation, quantum spin ice, artificial spin ice, and so on—are already implicit in this figure (though it took a lot of hard work to elucidate them!). In the abstract, we indicated that our aim has been to search out the biggest insight to have emerged from more than two decades of the study of spin ice. We are now in a position to decide what it is. If one looks back at the massive literature on vertex models one can see how one simple concept introduced by Pauling—that of weighted vertices—already generated an enormous body of interesting and valuable physics. Usually, when one modifies a simple model to make it more realistic for a particular case, it gets much more complicated and less generic, without getting much more interesting. Spin ice is the striking exception to this rule. Most unexpectedly, the addition of a couple of the more realistic aspects of water ice to the model—the dipole interactions and quantum fluctuations—retains the simplicity and broad relevance of the original model, including all the vertex physics, while generating a huge and rich landscape of new phenomena and concepts that will be explored and studied for years to come. This we see as the biggest insight to have emerged from all the research on spin ice: the massive and unexpected expansion of the scope of simple vertex models to describe diverse real systems, real phenomena and exotic physics. The increasing number of publications that refer to ‘spin ice’ (figure 9) is surely testament to this 17.

Appendix A. Summary of the magnetism of rare earth insulators

Here we outline the key features of magnetism in rare earth insulators, particularly as regards the realisation of simple spin models.

The elements of the rare earth or lanthanide series generally form trivalent (3+) ions with partly filled 4f-shells. As regards the free ions, electron repulsion of order 10^4 K and spin orbit coupling of order several thousand kelvins conspire to create a discrete series of single-ion states. These may be classified in the Russell–Saunders coupling scheme by the quantum numbers S, L, J of, respectively, the summed spin, orbital and total (spin plus orbital) angular momenta of the electrons. Owing to the large energy scales involved, it is often a good approximation to assume that only the ground term (a single J), given by Hund’s rules, is relevant to magnetic properties. For, example the free ion ground state of Ho^{3+} with 10f-electrons is labelled 5/2, which translates to S, L, J = 2, 6, 8 respectively while that of Dy^{3+} with 10f-electrons is labelled 7/2, which translates to S, L, J = 3, 5, 6 respectively.

17 To reduce such a mass of science to a short review has inevitably meant we have been forced to leave out, or given scant attention, to some excellent papers: the review is not meant to be comprehensive and we apologise to any authors who may find their work under-represented.
9f-electrons is $^6H_{15/2}$ which translates to $S, L, J = 5/2, 5, 15/2$ respectively.

A single ion ground state defined by $J$ consists of $2J + 1$ degenerate sublevels with quantum numbers $m_J = J, J - 1, \ldots, -J$. In a magnetic field $B$, these states split according to the first order Zeeman effect, $E = -g_J \mu_B m_J$ (where $g_J$ is the Landé $g$-factor), creating an evenly-spaced ladder of states. However, in a crystal, the $m_J$ states are already split and mixed together by the crystal electric field of surrounding negative ions. In many cases, a series of doublet states is produced, though more complex splitting schemes are possible. Kramers’ theorem shows that singlets can only occur in ions with an even number of f-electrons—all the states of Dy$^{3+}$, for example, are at least 'Kramers doublets'.

In general the resulting crystal field split states of the ion in a crystal consist of admixtures of the $m_J$ states. For example, one could have a doublet with a subtitle of the form:

$$\psi = a|J\rangle + b|J - 1\rangle + c|J - 2\rangle + \ldots,$$

and complex coefficients $a, b, c$ etc. The doublet Zeeman splits linearly in a magnetic field and can be represented by an effective spin one doublet half doublet with Zeeman energy $E = -g_{\text{eff}} \mu_B m_J$, where $m_J = \pm 1/2$ and the effective $g$-value $g_{\text{eff}}$ is $g_J$ multiplied by a function of the $a, b, c$ etc. The particular values of these coefficients depend on the symmetry of crystal field and the wavefunctions of the particular rare earth ions: several of them may be identically zero, or simply very small. It is this diversity of possible admixtures that makes each rare earth ion, in each crystal environment, very individual as regards its magnetism. In particular for large $J$, effective doublets dominated by small $m_J$ have very large XY-like quantum fluctuations with respect to a quantization axis, as occurs in Er$_2$Ti$_2$O$_7$ for example [121], while those dominated by large $m_J$ may approach classical Ising-like behaviour. This is what happens in Ho$_2$Ti$_2$O$_7$ and Dy$_2$Ti$_2$O$_7$, where the rare earth states approach pure Ising-like doublets with maximum $m_J = \pm J$, i.e. $\psi \approx |\pm J\rangle$. In this case the ionic magnetic moments are simply $\mu = g_J \mu_B \approx 10 \mu_B$ in both cases, and hence particularly large. This is an extreme type of behaviour: many other types of behaviour are possible including 'non magnetic' (singlet) or quadrupolar ground states, or even 'non-magnetic doublets' (those with vanishing first order Zeeman effect).

The concept of a $g$-tensor applies to any magnetic ion, but it is especially important in rare earth magnetism because the effective spins may be strongly quantised with respect to a local crystallographic axis, $z$, that may be different to the axis of the applied field. In that case the Zeeman energy is $E = -\mu_B \sum_{n\neq 0} f^{\text{eff}} B^n s^z$ and the principal axes of the $g$-tensor are determined by the crystallographic axes that define the crystal field. For the case of a single doublet, as discussed above, the Zeeman operator is diagonal in the $|m_J\rangle$ basis only when the field is applied parallel to the crystal field axis $z$ and then $g_{\text{eff}} = g_J$ corresponds to $g^2$. The $g$-tensor anisotropies in rare earth salts can be very large, as spin ice illustrates: in Ho$_2$Ti$_2$O$_7$ and Dy$_2$Ti$_2$O$_7$, the $g$-tensor referred to its principal axes is diagonal with $g_{∥} \approx 20$ and $g_{\perp} \approx 0$ to a close approximation [24]. This generates the moments $\mu = g/(1/2)\mu_B \approx 10 \mu_B$ mentioned above.

The principal interest in the field of ‘model magnetism’ is the effect of coupling between paramagnetic rare earth ions. Relative to the d-electrons of transition metal ions like Fe$^{3+}$, Co$^{2+}$ etc, the magnetic f-electron shell of the rare earth ions is tightly bound into the core. This has the consequence that there is little overlap of wavefunctions with neighbouring ions or ligands in a crystal, so superexchange interactions are very weak, typically of order 1 K or less. However, owing to the possibility of large moments, as discussed above, the dipole–dipole interaction between two ions can be relatively strong—of several kelvins, and as a long range interaction, this may be strongly amplified by collective effects, as it does in spin ice.

The rare earth ions thus present several differences to the perhaps more familiar magnetism of the d-block ions. The rich complexity of d-block magnetism arises because crystal fields, spin orbit coupling (and sometimes superexchange) are of similar magnitude. In the rare earth series, things are much more simple, in that spin orbit coupling, crystal field and superexchange (or dipole–dipole coupling) are successive perturbations of diminishing strength. But this simplicity does not imply dullness: rare earth ions may be regarded as elementary quantum-mechanical entities of considerable character that may be arranged in crystals in diverse ways.

From the model magnetism perspective, rare earth salts offer a number of advantages over their d-block counterparts. First, the fact that the f-electrons are relatively deep in the core, means that the chemistry of rare earth salts with different f-block ions tends to be similar, so typically one has access to a series of isomorphic crystal structures, in which interactions and quantum effects may be tuned by changing the rare earth ion. Second, rare earth salts are typically in the localised electron limit, which simplifies theoretical interpretations. Third, the weakness of the inter-ionic interactions means that in most cases, the whole of the interesting part of the field-temperature phase diagram is available to experiment, which is not always the case for d-electron systems. Pitted against this, the small energy scales mean that dynamics in interacting rare earth salts may be relatively hard to resolve by neutron scattering.

Spin ice exemplifies many basic aspects of rare earth magnetism rather well. The cubic pyrochlore materials Ho$_2$Ti$_2$O$_7$ and Dy$_2$Ti$_2$O$_7$ are simple, insulating ionic salts, whose paramagnetism arises from Ho$^{3+}$ and Dy$^{3+}$ respectively (Ti$^{4+}$ and O$^{2-}$ having full shells). The magnetic rare earth ions are subject to trigonal crystal fields with a strong axial component, such that the ground state doublet is $|\pm J\rangle$ to a close approximation. This means that the $g$-tensor has only one finite component, $g_{zz} \approx 20$, which creates two state Ising-like moments, locked along local anisotropy axes. There are four such axes in the pyrochlore structure (see figure 1). The relative complication of having four local axes, coupled with a strong long-ranged and anisotropic dipole–dipole interaction, arising from the large moments (10 $\mu_B$), makes these materials at first sight look like rather complicated dipolar magnets, and early studies noted them as such [14]. To understand why the actual behaviour of these materials is in many ways very simple, yet
full of rich possibilities, is a rather complex challenge that is addressed in the main text.

**ORCID iDs**

Steven T Bramwell [ORCID: 0000-0002-4097-5465]

Mark J Harris [ORCID: 0000-0002-7386-7743]

**References**

[1] Pauling L 1935 *J. Am. Chem. Soc.* **57** 2680
[2] Bernal J D and Fowler R H 1933 *J. Chem. Phys.* 1 515–48
[3] Giauque W F and Stout J W 1936 *J. Am. Chem. Soc.* **58** 1144
[4] Slater J C 1941 *J. Chem. Phys.* 9 16
[5] Lieb E H 1967 *Phys. Rev.* 162 162
[6] Baxter R J 1982 *Exactly Solved Models in Statistical Mechanics* (New York: Academic)
[7] Anderson P W 1956 *Phys. Rev.* **102** 1008
[8] Reimers J N, Greedan J E, Stager C V, Björgvinnsen M and Zarestky J L 1969 *J. Appl. Phys.* 40 4603
[9] Bovo L, Moya X, Prabhakaran D, Soh Y-A, Boothroyd A, Coull C, Moessner R, Czternasty C, Meissner M, Rule KC, Hoffmann J B and Gardner J S 2009 *Nat. Phys.* 5 258
[10] Barry K, Zhang B, Anand N, Xin Y, Vailionis A, Neu J, Heikes C, Cochran C, Zhou H, Qiu Y, Ratcliff W, Siegrist T and Beckman C 2019 *Phys. Rev. Mater.* 3 084412
[11] Henley C L 2005 *Phys. Rev. Lett.* 95 011007
[12] Jaubert L D C and Holdsworth P C W 2014 *Nat. Phys.* 10 1219
[13] Zhou H D, Bramwell S T, Cheng J G, Wiebe C R, Li G, Balicas L, Bloksom J A, Silverstein H J, Zhou J S, Goodenough J B and Gardner J S 2011 *Nat. Commun.* 2 478
[14] Brooks-Bartlett M E, Banks S T, Jaubert L D, Harman-Clarke A and Holdsworth P C W 2014 *Phys. Rev. X* 4 011007
[15] Jaubert L D C, Chalker J T, Holdsworth P C W and Moessner R 2008 *Phys. Rev. Lett.* **100** 067207
[16] Matsuhira K, Hiroi Z, Tayama T, Takagi S and Sakakibara T 2002 *J. Phys.: Condens. Matter* **14** L559
[17] Aoki H, Sakakibara T, Matsuhira K and Hiroi Z 2004 *J. Phys. Soc. Jpn.* 73 2851
[18] Sakakibara T, Tayama T, Hiroi Z, Matsuhira K and Takagi S 2003 *Phys. Rev. Lett.* 90 207205
[19] Moessner R and Sandhi S L 2003 *Phys. Rev. B* 68 064411
[20] Fennell T, Bramwell S T, Mcmorrow D F, Manu P and Wildes A R 2007 *Nat. Phys.* 3 566
[21] Hiroi Z, Matsuhira K and Ogata M 2003 *Phys. Soc. Jpn.* 72 3045
[22] Zhou H D, Cheng J G, Hallas A M, Wiebe C R, Li G, Balicas L, Zhou J S, Goodenough J B, Gardner J S and Choi E S 2012 *Phys. Rev. Lett.* **108** 207206
[23] Lhotel E, Petit S, Guittény S, Florea O, Ciomaga Hatnean M, Colin C, Ressouche E, Lees M R and Balakrishnan G 2015 *Phys. Rev. Lett.* 115 197202
[24] Matrak P, Coneri F, Hoek M, Turner S, Idrissi H, Van Tendeloo G and Hilgenkamp H 2014 *Apl. Mater.* 2 032101
[25] Barry K, Zhang B, Anand N, Xin Y, Vailionis A, Neu J, Heikes C, Cochran C, Zhou H, Qiu Y, Ratcliff W, Siegrist T and Beckman C 2019 *Phys. Rev. Mater.* 3 084412
[26] Vedmedenko E Y 2016 *Phys. Rev. Lett.* 116 077202
[27] Henley C L 2005 *Phys. Rev. Lett.* 91 014424
[28] Castelnovo C, Moessner R and Sandhi S L 2008 *Nature* 451 42
[29] Ryzhkin I A 2005 *J. Exp. Theor. Phys.* 101 481
[30] Jaccard C 1964 *Phys. kondens. Mater.* 3 99
[31] Ryzhkin M I, Ryzhkin I A and Bramwell S T 2013 *Europhys. Lett.* 104 37005
[32] Jaubert L D C and Holdsworth P C W 2009 *Nat. Phys.* 5 258
[33] Morris D J P, Tennant D A, Grigera S A, Klemke B, Castelnovo C, Moessner R, Czternasty C, Meissner M, Rule K C, Hoffmann J U, Kiefer K, Gericke S, Slobinsky D and Perry R S 2009 *Science* 326 411
[34] Kadowaki H, Doi N, Aoki Y, Tabata Y, Sato T J, Lynn J W, Matsuhira K and Hiroi Z 2009 *J. Phys. Soc. Jpn.* 78 103706
[35] Bramwell S T, Giblin S R, Calder S, Aldus R, Prabhakaran D and Fennell T 2009 *Nature* 461 956
[36] Dumsiger et al 2011 *Phys. Rev. Lett.* 107 207207
