Pergamon

Stud. Hist. Phil. Mod. Phys., Vol. 30, No. 1, pp. 85–110, 1999

Printed in Great Britain
1355-2198/99 $ – see front matter

$g-2$ and the Trust in Experimental Results

***B. Lautrup*† *and H. Zinkernagel*‡**

1. Introduction

There was a time when it was often held that it is the reproducibility of experiments which establishes experimental results as objective facts. In the wake of Kuhn, however, it was argued that theoretical presuppositions shape or even determine experimental results. And since theories change, so will the results of experiments. Consequently the 'objectiveness' of experimental results became relative to their theoretical framework. To be sure, there have been numerous objections to just how radical theory changes were and thus how different experimental results could be in different theories. In any case, this Kuhn-inspired emphasis on theories has been opposed by a recent philosophy of experiments which has argued that experiments can remain stable when theories change — experiments have a life of their own (see for example Hacking, 1983; Franklin, 1986 and Galison, 1987).

Have experimental results then again become objective facts about nature? Answers differ, but it should be stressed that the philosophy of experiments has not renounced theory as such. Rather the relation between theory and experiment has been seen in a new light. For instance, even though theories do not determine experimental results, some theory or background suppositions are needed in order to make sense of experiments. Building on such insights the question has been raised of how experiments end. When is the scientific community prepared to believe in an experimental result? This way of putting the

(*Received 18 February 1998; revised 14 September 1998*)
† The Niels Bohr Institute, University of Copenhagen, Denmark (*e-mail*: lautrup@nbi.dk).
‡ Center for the Philosophy of Nature and Science Studies, The Niels Bohr Institute, University of Copenhagen, Denmark (*e-mail*: zink@nbi.dk).

PII: S1355-2198(98)00032-X

question, however, assumes that experiments *do* end. Which of course they often do. But some experiments, or rather experimental studies of the same questions, are repeated again and again.

In this paper we want to reflect on the development since 1947 of experiments on the magnetic moment of the electron, commonly referred to as $g-2$ [g minus 2] experiments. The ancestors of these experiments were the gyromagnetic experiments which have been chronicled by Galison (1987) in his book *How Experiments End*. Galison provides an analysis of gyromagnetic experiments from 1913 until around 1933 and discusses how the experiments survived through major theory changes. The period covered by Galison is spanned by classical electromagnetism, the old and new quantum mechanics and relativistic quantum mechanics. But experiments on the magnetic properties of electrons did not end with Galison's analysis. In fact, the continuing series of experiments on the magnetic moment of the free electron covered in this article provides the most accurate test of Quantum Electrodynamics (QED), and refinements continue to this day.[1] Nowhere else in physics has a theory been confronted with experimental results to such high accuracy.

It is sometimes assumed that repetitions of experiments only take place in areas of controversy, for instance to test the stability of a new effect under variation of the experimental circumstances (see e.g. Collins, 1984). The $g-2$ experiments have all been performed in a period under a fixed theoretical framework, QED. Nevertheless, the development of these experiments provides an interesting example of the interplay between theory and experiment. As we shall see, the $g-2$ experiments appear well suited for a discussion of questions raised by recent philosophy and history of science, for instance regarding some of the elements contributing to the trust in experimental results.

Our point of departure will be the concept of errors which couples nicely to the debate about theory-ladenness of experiments. At every point in the history of the $g-2$ experiments, a certain amount of theory was necessary to convert the raw measurements into a value for $g-2$. To discuss what theoretical considerations were involved in the various $g-2$ experiments, we explain below some of the ideas on which the experiments were built.

Concerning our case study, it is important to stress that although we do include the scientists' published motivations for their work, we do not undertake a detailed historical analysis of the particular circumstances (e.g. economic, social, personal, etc.), leading to the published articles.[2] Instead we shall attempt to extract philosophical lessons from the published articles by seeing them in relation to the historical development of the experiments.

[1] Some remarks about the experimental situation in the period 1933–1947 may be found in Schweber (1994).

[2] By this we do not mean to neglect the value of such studies. The importance of going behind the reconstructed logical ordering of the published papers has been emphasised by many recent scholars. See for instance (Collins, 1984) and (Galison, 1987, p. 244).

Thus, rather than asking how experiments end, we will be asking why experiments *continue*.

2. Errors and Error Bars

Experiments are beset with errors. As has been pointed out before in the literature, a significant and constitutive part of experimental practice is estimations of, and corrections for, errors — extracting signal from noise or foreground from background. Since a central emphasis in this article is on experimental errors it is appropriate to give a short introduction to the concepts of statistical and systematical errors.

Statistical errors are random errors. They arise from the finite accuracy of the measuring and monitoring apparatus or inherent randomness of the phenomena under scrutiny, and lead to a spread of the experimental result around an average value. The statistical errors are assumed truly random, so the size of the statistical error in one measurement is independent of errors in other measurements of the same quantity. Ideally, if there were no other errors than statistical errors, the average value taken from an indefinite number of measurements would constitute the 'true' value for the measured quantity. Statistics deals with these errors by taking into account that a quantity can only be determined a finite number of times.

But experimental results are also subject to systematical errors. These arise from experimental effects not taken into account and/or bias in the extraction of results from data. In contrast to the statistical errors, a systematical error does not imply a spread around a central value but merely shifts the result away from the true value.

It is common to state an experimental result with error bars. For instance, if one has measured a physical quantity, the result x_{exp} can be reported as $x_{\mathrm{exp}} \pm \Delta$ where Δ indicates the error. If the error is expected to be mainly statistical, this is usually taken to mean that the interval $[x_{\mathrm{exp}} - \Delta; x_{\mathrm{exp}} + \Delta]$ includes the true value x_{true} with 68.3% probability (see e.g. Particle Data Group, 1994, p. 1278). Accordingly, it will not raise concerns if a later experiment gives x'_{exp} with error Δ' and the corresponding interval is slightly outside the interval of x_{exp} and Δ. In the $g - 2$ experiment we shall see that the systematical errors have often been the major source of error and hence that discrepancies between succeeding measurements and their error bars have been regarded as problematic.

The systematic errors can be divided into three groups. First, there are known systematic effects which can be corrected for either experimentally or theoretically (and for this reason these effects are called corrections). In the final reported result of the measurement, corrections will be incorporated directly in the experimental value and are not reflected in the error bars. The second type of systematic errors refers to effects which are thought or known to play a role but whose exact influence cannot be determined. In such cases the error may also be estimated either by theoretical or experimental arguments. It may happen, however, that an effect is thought or known to play a role but that for some reason it cannot be theoretically or experimentally estimated, in which case the estimate may be a more or less educated guess. In any case, if an estimate of a systematic effect is made, it is reflected in the error bars. Finally, the third category of systematic errors consists of those errors which are unknown at the time of the measurement and consequently cannot be corrected or shown in the error bars.

Thus, theory will dictate what systematic errors can be expected in an experiment, and theory may be used to estimate their influence. In the account of the $g-2$ experiments below we shall pay attention to what kind of theoretical arguments were involved in the process of extracting a value for $g-2$ from the measurements. Before turning to our case study we provide some background on the origin of magnetism and the g-factor (see Galison (1987) for a full account).

3. The *g*-factor

In Maxwell–Lorentz electrodynamics there are no elementary magnetic charges. Although the equations leave room for the possibility, it seems now, after many years of fruitless experiments, that free magnetic charges in the form of lonely north poles or south poles (monopoles) indeed do not exist (see e.g. Klapdor–Kleingrothaus and Staudt, 1995).

From the assumption that the electron is the only charge carrier responsible for magnetism and the fact that all electrons have the same charge and mass follows the prediction that the density of charge should be proportional to the density of mass for these charge carriers. This proportionality leads in turn to a relation between the magnetic moment of a current distribution, which is a measure of its magnetic field, and the angular momentum of its mass distribution, which is a measure of its state of rotation. They must in fact be proportional to each other with a constant of proportionality given by the ratio of the electron's charge to twice its mass ($e/2m$).

If the electron were not the only charge carrier things would be different. Contamination from another charge carrier with a different ratio between charge and mass would lead to a different constant of proportionality. In order to include this possibility a 'fudge-factor' g was introduced (by Landé in 1921 (Galison, 1987, p. 64)) to take care of such deviations (so the ratio was written $g\,e/2m$). This g-factor or gyromagnetic ratio would accordingly be exactly one (i.e. $g=1$) if the electron were the only charge carrier.

The curious history of the gyromagnetic experiments illustrates the importance of theoretical ideas for the outcome of experiments. Around 1915 the prejudice of the theorists (among them Einstein) was strongly in favour of $g=1$ and experimental results were also found in this neighbourhood at that time, and as late as 1923 by Einstein's collaborators. Other experimentalists were also influenced but did eventually abandon their theoretical prejudices. It took nevertheless the concerted efforts of many experiments to indisputably dislodge the measured value from the expected one. Galison concludes in his analysis of the background of this situation that the theoretical prejudice did not by itself bias the experimental result, but can possibly have created a mindset in which experiments were terminated and the search for systematic errors given up when a result was found near the strongly expected one (Galison, 1987).

In 1925, Goudsmit and Uhlenbeck (1926) explained the so-called anomalous Zeeman effect (see below) by introducing the electron spin, a quantum phenomenon akin to an internal rotation of the electron about an axis. Using a gyromagnetic factor of $g = 2$ for the magnetic moment associated with the spin, they were able to explain the fine structure doubling of spectral lines and the Zeeman effect. Thus, magnetism in materials turned out to be more complicated than previously thought, being a mixture of contributions due to orbital electron motion with $g = 1$ and the intrinsic electron spin with $g = 2$. In 1928 Dirac published his relativistic theory in which the electron is born with a spin and a gyromagnetic factor exactly equal to 2. For the next two decades this became a theoretical prejudice which agreed comfortably with experimental results (see Schweber, 1994, pp. 211ff).

4. Modern Experiments on the Electron g-factor

The first suggestion that the g-factor of the electron might be *different* from 2 was made by Breit (1947) (see also Combley, 1979 and Schweber, 1994, p. 220), and was prompted by a disagreement between theory and precise measurements of the hyperfine structure of hydrogen obtained by Nafe, Nelson and Rabi (1947).

This began a series of experiments for determining the precise value of the difference between the actual g-factor of the electron and the Dirac value 2. The overarching motivation for the experiments (as well as the theoretical calculations) was to determine this difference to an ever greater precision. One may roughly divide the modern development into three different phases that more or less follow each other sequentially in time: (1) atomic level experiments, (2) free electron spin precession experiments, and (3) free electron spin resonance experiments. We shall discuss these in turn below.

In Table 1 the experimental determinations of the g-factor of the electron or rather the corresponding anomaly $a = (g - 2)/2$ are listed. The same data are also plotted in Fig. 1, but because of the rapid drop in the size of error bars (see Fig. 2) the plot is not representative of the two later phases. See Fig. 2 for another presentation of the data. It should be noted that not all experimental values refer to independent measurements.

The decrease in experimental errors over the years has been remarkable. As shown in Fig. 2, the decreasing errors roughly follow an exponential curve from 1958 to 1984. On average the error decreases by a factor of 1.8 per year during this twentysix-year period.

Table 1. *Experimental determinations of the electron g-factor anomaly* ($a = (g-2)/2$). *The error is written in the parenthesis after the value and refers to the last digits. The square boxes represent atomic level experiments, the stars free electron spin precession experiments, and the diamonds free electron spin resonance experiments. A bullet indicates a re-evaluation of earlier experiments with no new data taken. The theoretical values are only included when they change and are the actual values quoted by the experimentalists in the comparison of theory and experiment. Notice that all anomalies have been multiplied by* 10^3.

	Authors	Year	Type	Experiment $\times 10^3$	Theory $\times 10^3$
	Kusch, Foley	1947	■	1.15(4)	
	Foley, Kusch	1948	■	1.22(3)	
	Kusch, Foley	1948	■	1.19(5)	1.162
	Koenig, Prodell, Kusch	1952	■	1.146(12)	1.145,4
	Beringer, Heald	1954	■	1.148(6)	
	Louisell, Pidd, Crane	1954	★	0(5)	
	Franken, Liebes	1956	■	1.165(11)	
	Dehmelt	1958	◆	1.116(40)	
	Schupp, Pidd, Crane	1961	★	1.160,9(24)	1.159,6
	Farago, Gardiner, Muir, Rae	1963	★	1.153(23)	
	Wilkinson,Crane	1963	★	1.159,622(27)	1.159,615
•	Rich	1968	★	1.159,557(30)	1.159,617
•	Farley	1968	★	1.159,596(22)	
	Graff, Major, Roeder, Werth	1968	◆	1.159(2)	
	Graff, Klempt, Werth	1969	◆	1.159,66(30)	
•	Henry, Silver	1969	★	1.159,549(30)	1.159,641
	Wesley, Rich	1970	★	1.159,644(7)	1.159,644
	Wesley, Rich	1971	★	1.159,657,7(35)	1.159,655(2)
•	Granger, Ford	1972	★	1.159,656,7(35)	1.159,655(2)
	Walls, Stein	1973	◆	1.159,667(24)	1.159,652,9(24)
	Van Dyck, Schwinberg, Dehmelt	1977	◆	1.159,652,410(200)	
	Van Dyck, Schwinberg, Dehmelt	1979	◆	1.159,652,200(40)	1.159,652,34(31)
	Van Dyck, Schwinberg, Dehmelt	1984	◆	1.159,652,193(4)	1.159,652,460(145)
	Van Dyck, Schwinberg, Dehmelt	1987	◆	1.159,652,188,4(43)	1.159,652,263(104)
•	Van Dyck	1990	◆	1.159,652,189(4)	1.159,652,133(29)

Apart from the very first experiment (Kusch and Foley, 1947) and the most recent series of experiments (Van Dyck, Schwinberg and Dehmelt, 1977), the theoretical value for the electron g-factor was always known to higher precision than the experimental values. The theoretical predictions changed due to more precise QED calculations (Lautrup, Peterman and de Rafael, 1972; Kinoshita, 1995) and to changes in the measured values for the fine structure constant α (see below). We shall discuss the theoretical calculations of $g-2$ only insofar as they are directly related to the experiments. In the following we describe the physical principles behind the experiments in order to highlight some of the systematic corrections applied to the raw data when obtaining the final quoted experimental results. Our account below is by no means exhaustive but covers mainly those features of the experiments that are relevant to our discussion.

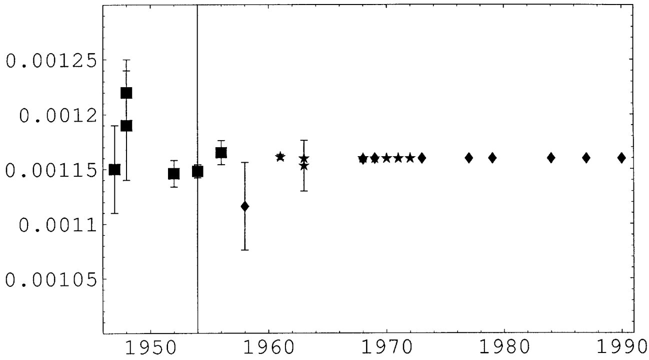

Fig. 1. *Modern measurements from 1947 to 1987 in terms of the anomaly* $a = (g - 2)/2$. *The square boxes represent atomic level experiments, the stars free electron spin precession experiments, and the diamonds free electron spin resonance experiments. Error bars are included everywhere, but are too small to be seen in the plot after 1964. The vertical line in 1954 is part of the large error bar from the pilot experiment on spin precession by Louisell, Pidd and Crane* (*1954*). *See also* Table 1.

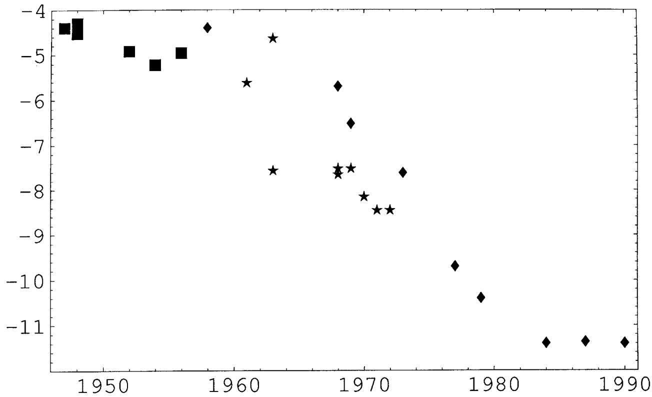

Fig. 2. *Logarithm of experimental errors plotted versus year. The square boxes represent atomic level experiments, the stars free electron spin precession experiments, and the diamonds free electron spin resonance experiments. Note that the precision of the spin precession experiment by Louisell is too low to be seen in the figure* (*see* Table 1).

5. Atomic Level Experiments

Stimulated by suggestions by Breit and Rabi, Kusch and Foley in 1947 carried out high-precision measurements of atomic levels revealing a discrepancy which might be due to an anomalous g-factor.

The experiments were based on the so-called Zeeman effect which denotes the splitting of atomic levels into sublevels in a magnetic field. The effect is caused by

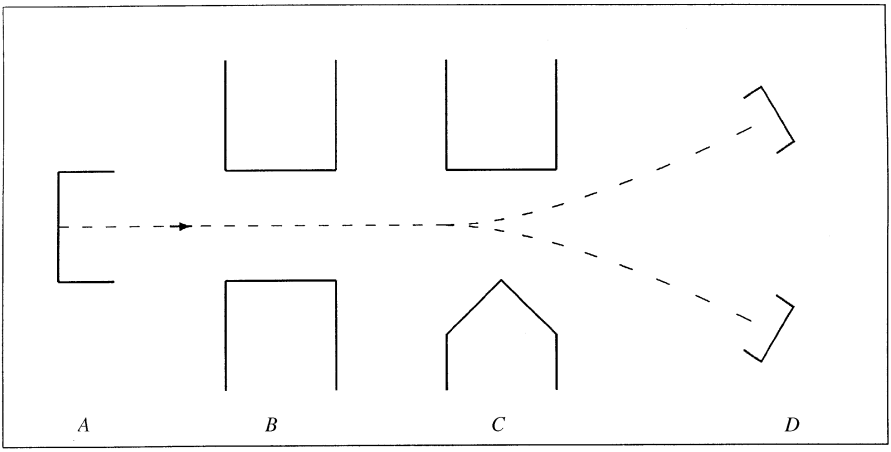

Fig. 3. *Sketch of atomic level experiments for the determination of the atomic g-factor. A beam of atoms are produced (A) and sent through a strong homogeneous magnetic field (B) with a superposed weak oscillating magnetic field, causing transitions between the Zeeman levels of the atoms. An inhomogeneous magnetic field (C) afterwards splits the beam into sub-beams (here two) according to the Zeeman levels, just as in the Stern–Gerlach experiment. The sub-beam intensities are finally monitored (D) as a function of the frequency of the oscillating magnetic field, allowing determination of the transition frequency from the position of the peak in intensity at resonance.*

interaction between the magnetic field and the total magnetic moment of the atom and each sublevel corresponds to a different orientation of the magnetic moment. The actual measurement consisted in subjecting a beam of atoms to a weak oscillating magnetic field on top of a strong homogeneous one, and determining the frequency required to excite transitions between the Zeeman levels. The state of the atoms after the excitation was observed by splitting the beam into sub-beams corresponding to the different Zeeman levels (a Stern–Gerlach type set-up, see Fig. 3). Sweeping the oscillation frequency across the natural transition frequency of the atoms, a marked peak could be observed in one of these sub-beams.

According to quantum mechanics the atomic transition frequency is

$$\omega_A = g_A \frac{e}{2m} B, \tag{1}$$

where $e/2m$ is the Bohr magneton, B is the magnetic field and g_A (in analogy with the electron g-factor) is the magnetic moment of the atom in units of the Bohr magneton.[3] The atomic g-factor is partly due to the g-factor for orbital motion

[3] The Bohr magneton is given by $e\hbar/2m$ where e is the magnitude of the electron's charge, m its mass and $\hbar = h/2\pi$ is Planck's reduced constant. The g-factor of the electron is simply its magnetic moment in units of the Bohr magneton.

of electrons around the nucleus and partly to the g-factor for the intrinsic spin of the electrons. In such experiments it is only possible to determine the combined effect of these two contributions, not each individually.

It is technically difficult to obtain a sufficiently precise value for the magnetic field strength. This problem disappears if one calculates the ratio of transition frequencies for two different transitions, 1 and 2, in the same or in different atoms,

$$\frac{g_1}{g_2} = \frac{\omega_1}{\omega_2}. \tag{2}$$

From such ratios of atomic g-factors the ratio of the spin and orbital g-factors could be extracted.

As we mentioned above, the theoretical expectation after Dirac's relativistic theory of the electron was that orbital motion was associated with $g = 1$ and spin motion with $g = 2$. In their first paper, Kusch and Foley (1947) found a discrepancy with these assumptions and noted that it could be corrected by adding an anomalous contribution to the g-factor of either the orbital motion or the spin. In their second paper, Foley and Kusch (1948) abandoned the first possibility and in a footnote quoted Schwinger for the theoretical justification (see below). The anomaly was consequently understood as entirely due to an *anomalous* magnetic moment for the electron. With this interpretation they found the final value[4] of the electron g-factor anomaly to be $a = 0.00119(5)$ (Kusch and Foley, 1948).

From this moment onwards the experimental focus was on the discrepancy between the Dirac value 2 and the actual value of g. It is now customary to quote all experimental as well as theoretical results in terms of the electron anomaly which as mentioned above is half the difference between the actual value and the Dirac value (see Table 1).

The first theoretical calculation of the free (i.e. not bound to any system) electron's anomaly by Schwinger in 1948 gave the result $a = \alpha/2\pi \simeq 0.001162$, where $\alpha \approx 1/137$ is the dimensionless fine structure constant known from atomic physics. The fine agreement between theory and experiment was perceived as a confirmation of the internal radiative processes involving a single photon predicted by QED.[5] Later theoretical efforts have entirely been concerned with the calculation of higher order radiative effects involving more than one internal photon, corresponding to higher powers of α.

Foley and Kusch (1948) explicitly pointed out that the theoretical estimate by Breit (1947), who among others had inspired them to do the experiment, was in disagreement with their experimental result.[6] But they were, as mentioned,

[4] This value is actually an average over three experimental runs using different choices of atoms and levels.

[5] The difficulties in calculating radiative corrections from QED arose from mathematical infinities which were first circumvented with the renormalisation scheme of Schwinger, Feynman and Tomonaga in and around 1947 (see e.g. Schweber, 1994).

[6] See Schweber (1994, p. 222) for an account of Breit's discontent with Foley and Kusch's remarks.

aware of Schwinger's work in progress, as he was of theirs (Schwinger, 1948). From the published papers (Foley and Kusch, 1948; Schwinger, 1948) it is not clear whether they knew Schwinger's specific value for a, which was in perfect agreement with their measurement (in his paper Schwinger pointed out that Breit had not done the calculation correctly). In any case, it clearly took QED to associate the experimental result with a deviation of the spin g-factor rather than the orbital g-factor.

In the decade following the pioneering experiments by Kusch and Foley similar experiments were carried out with somewhat improved precision (see Table 1). A serious problem with atomic determinations of the g-factor for the electron arises, however, from the complicated corrections due to the electron being bound in an atom and not being free. At the level of the Schwinger calculation these atomic corrections may be ignored, but at the next level (of order α^2), where radiative processes involving two internal photons come in, it is necessary for the comparison with theory to include atomic (non-QED) corrections to the same precision.

In 1952, Koenig, Prodell, and Kusch indeed applied a relativistic mass correction of order α^2 to their measurement of the ratio between the g-factor of the bound electron in hydrogen and the proton g-factor. They measured this ratio with an uncertainty which was many times smaller than the relativistic correction itself. Using the much less precise result for the proton g-factor obtained by Gardner and Purcell (1949), they finally arrived at a value for the electron g-factor with an uncertainty of the same order of magnitude as the relativistic mass correction and the second order QED correction by Karplus and Kroll (1950).[7] They obtained excellent agreement (experiment $a = 0.001146(12)$, theory $a = 0.0011454$). Beringer and Heald (1954) carried out a slightly modified experiment in 1954 and obtained a result which was in good agreement with Koenig *et al.* At this point in time the agreement between theory and experiment seemed perfect, even if unbeknownst to everybody the theoretical value was in error.

In the experiments following Kusch and Foley's the actually measured quantities are also transition frequencies. The experimental set-up with hydrogen beams only permits the determination of one transition frequency, ω_H, with the second transition frequency, ω_P, being obtained from nuclear spin resonance on protons in the same magnetic field. The g-factor of hydrogen is then determined by

$$g_H = g_P \frac{\omega_H}{\omega_P}, \tag{3}$$

[7] One should note that the measurement by Koenig *et al.* does not depend on the application of the relativistic correction. They measured the g-factor of the electron (in units of that of the proton) for the bound electron in hydrogen, and that result stands by itself without the relativistic correction. The relativistic correction was entirely necessitated by the wish to produce a new value for the g-factor of the free electron, and the possibility of a comparison with theory.

where g_P is the proton magnetic moment in units of the Bohr magneton. Actually the right-hand side of this relationship does need systematic corrections mainly due to non-linearities in the atomic Zeeman levels as a function of the magnetic field, effects well-known at the time.

In order to determine the hydrogen g-factor g_H (which contains the electron g-factor) in these experiments, it is necessary to obtain a value for g_P. In both experiments this value was taken from an earlier experiment by Gardner and Purcell (1949) and the uncertainty in this quantity dominated the resulting uncertainty in g_H and thereby g.

In 1956, Franken and Liebes remeasured the proton magnetic moment in an experiment designed to eliminate the errors in the g_P determination due to the influence of non-vanishing electrostatic fields present in their apparatus. These errors were only estimated theoretically but not measured directly in the earlier experiment by Gardner and Purcell. The improvement was based on the idea that even if the actual electric field strength were unknown, its influence on the g-factor of the proton depended on the magnetic field strength and the electron velocity. Carrying out the experiments for a range of magnetic field strengths, Franken and Liebes were able to determine the size of the correction experimentally and subtract it from the measured values of g_P.

The new result for g_P disagreed with Gardner and Purcell's by about twice the quoted errors. Furthermore, in combination with the previous results by Koenig, Prodell and Kusch (1952), and by Beringer and Heald (1954), this measurement of g_P led to a new value for the g-factor in disagreement with the theoretical value of Karplus and Kroll by about twice the experimental error.

Franken and Liebes' experiment raised doubts about the agreement between theory and experiment. Even without the benefit of hindsight, which indicates that Gardner and Purcell must have underestimated their errors by at least a factor of two, the experiment of Franken and Liebes—in spite of the quoted uncertainties being the same as in the previous experiment—appears to be a better experiment, because they turned an educated guess by Gardner and Purcell concerning the electric field into an experimentally determined correction.

The theoretical value of Karplus and Kroll was found to be in error by Peterman (1957) and Sommerfield (1957). The experiment of Franken and Liebes served as an inspiration for the theoretical recalculations which again brought theory and experiment into agreement within about half the experimental error.

6. Free Electron Spin Precession Experiments

In the preceding experiments the g-factor was measured for electrons bound in atoms. The complicated corrections due to the binding influenced, as we have discussed, the interpretation of the atomic experiments. If on the other hand it were possible to determine the g-factor of the *free* electron, such corrections would be absent.

In 1954 a pilot experiment by Louisell, Pidd and Crane (1953, 1954) demonstrated the feasibility of a new method for determining the g-factor anomaly of the free electron. The central feature of the method consists in setting electrons in circular motion in a plane orthogonal to a homogenous magnetic field. The number of revolutions per second is measured by the angular velocity, ω_c, called the cyclotron frequency. The magnetic moment of the electron will simultaneously precess around the direction of the magnetic field and if the g-factor were exactly 2, its precession rate, or angular velocity ω_s, would be equal to the cyclotron frequency ω_c, implying that spin magnetic moment and velocity would rotate at the same rate and maintain a constant angular separation.

Conversely, if g is not exactly 2, the angle between the direction of the electron velocity and the direction of the spin magnetic moment will change with a rate given by the difference $\omega_a = \omega_s - \omega_c$, which is proportional to the anomaly $a = (g - 2)/2$. This means that the anomaly may be determined as a ratio between two frequencies

$$a = \frac{\omega_a}{\omega_0}, \tag{4}$$

where the frequency in the denominator is $\omega_0 = eB/m$ with e and m the charge and mass of the electron.[8] This frequency is a direct measure of the magnetic field strength B. As before there are systematic corrections to this simple relation (see below).

The actual experimental set-up of Louisell, Pidd and Crane is based on an observation by Mott (1929) that unpolarised electrons scattering off atomic nuclei will get their spins partially polarised (see Fig. 4). The polarised electrons are then allowed to circulate in a homogenous magnetic field for a certain number of revolutions. Finally the electrons are scattered once more and here it is utilised that the intensity of the scattered polarised electrons for a particular scattering angle will depend on the direction of polarisation (Mott, 1929). Thus by observing the intensity variation of the scattered electrons as a function of the scattering angle, Louisell *et al.* could determine the final spin direction.

Due to the anomaly being of the order of one part in a thousand it takes a thousand orbital revolutions in the magnetic field for the direction of the spin magnetic moment to make one revolution relative to the initial direction. Louisell, Pidd and Crane only studied the electrons for five cycles of revolution, corresponding to a change of just two degrees in the angular separation of spin and velocity. The experiment only allowed them to conclude that the spin and

[8] This equation is also valid in the relativistic case. For a non-relativistic electron the cyclotron frequency is equal to ω_0. For a full discussion of relativistic corrections see for example Combley (1979).

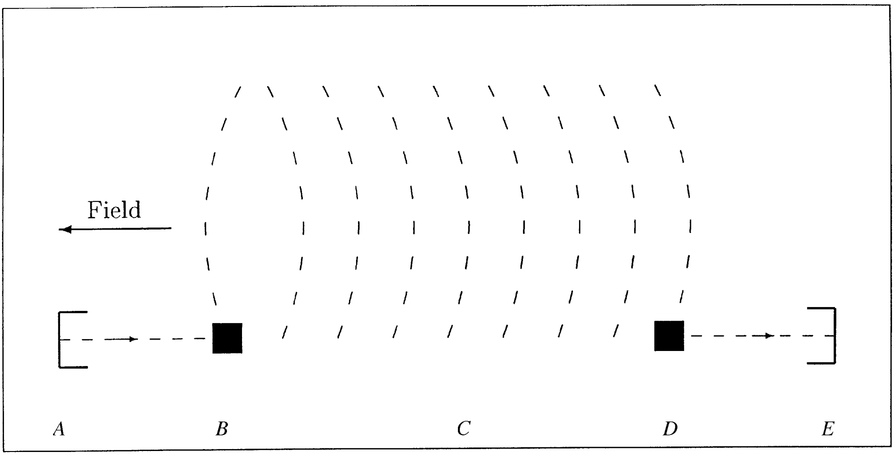

Fig. 4. *Sketch of free electron spin precession experiments for determination of the electron's g − 2. A bunch of electrons is produced (A) and (partially) polarised through 90° Mott scattering off the atomic nuclei of the target (B). The polarised electrons are trapped magnetically in a slightly inhomogeneous magnetic field for a large number of revolutions (C), and are finally allowed to hit another target (D), and once more get scattered through 90°. The counting rate for electrons arriving in the counter (E) varies cyclically as a function of trapping time with frequency ω_a which is a direct measure of the anomaly.*

velocity direction rotated at the same rate within the experimental resolution which was five times larger than the one necessary to observe the anomaly.

In 1961 Schupp, Pidd and Crane reported a highly improved version of this experiment in which the electrons were trapped for many thousands of revolutions instead of just five. By registering the actual trapping times for the polarised electrons they could determine the cyclic change in spin direction relative to velocity as a function of trapping time and thereby ω_a. The frequency $\omega_0 = eB/m$ was obtained from direct measurement of the magnetic field using Franken and Liebes' result. The final quoted result $a = 0.0011609(24)$ agreed with the theoretical calculation by Sommerfield (1957) and Peterman (1957), $a = 0.0011596$, to within half the experimental error.

The authors' own trust in their result is however somewhat less than complete. The experiment was carried out for several values of the magnetic field strength and the simple weighted average over all runs came to $a = 0.0011627$, with a statistical error of less than half a unit at the last digit. The distance between this result and theory is more than 60 times the statistical error. The largest estimated systematic error stems from inhomogeneities of the magnetic field necessary to create the magnetic trap. Put together with a number of smaller estimated systematic errors the authors in the end adopt an estimated total value for the systematic error of 14 units (on the last two digits). This brings the distance between theory and experiment down to about twice the experimental error.

In the experiment by Franken and Liebes (1956) discussed above, important systematic corrections due to stray electric fields could be eliminated by varying the magnetic field strength. Applying the same type of correction to their experiment, Schupp, Pidd and Crane were able to arrive at a measured value of $a = 0.0011609$, which brings theory and experiment into agreement within the above-mentioned error of 14 units on the last digit.

The authors are however not quite sure about this correction, in particular because it does not behave as expected under variation of some of the experimental conditions. They state that the correction is 'based on an uncertain hypothesis', namely that the dependency on the magnetic field strength is actually due to electric fields and not to some other instrumental cause, or even a real variation in the electron g-factor with the magnetic field (or equivalently, the electron velocity). The authors make the following comment about the use of this correction:

> In deciding upon a single value for a to give as the result of the experiment, our judgement is that we should recognize the trend [in the data corresponding to measurements with different magnetic fields], and proceed on the assumption that a radial electric field is present, in spite of certain weaknesses in the evidence for it.

In the end they published the value $a = 0.0011609$ but assigned to it an error which was great enough to include the weighted average over all measurements. The final published error thus became 24 on the last digits. This correction brought theory and experiment into agreement within half the experimental error.

In an experiment by Farago, Gardiner, Muir and Rae (1963) a transverse electric field was explicitly introduced to control the number of cyclotron revolutions. The experimental precision (see Table 1) was limited by unsurmountable technical problems (Rich and Wesley, 1972) and only attained the 1% level, but did otherwise agree with the previous experiments and theory.

In the same year an experiment by Wilkinson and Crane (1962, 1963) resulted in an order of magnitude increase in the precision obtained by Schupp, Pidd and Crane (1961). The experiment was an advance over the earlier one at several points. Central to the improvement was a reduction of the effects of stray electrostatic fields by increasing the separation between the electron orbits and the material of the vacuum chamber. The authors expressed no doubts about the need for the electric field correction which as before was applied after all other errors had been estimated.

This time the trend in the data was clear and eliminated the need for an *ad hoc* assignment of error. Instead the error was deduced from the estimated errors on the individual data points at different magnetic field strengths. The final published value for the anomaly became $a = 0.001159622(27)$. The theoretical value which still only included the two first radiative corrections was at this point in time $a = 0.001159615$. The agreement between theory and experiment was impressive, amounting only to one quarter of the experimental error. However, as we shall see below, this agreement was later to be cast into doubt.

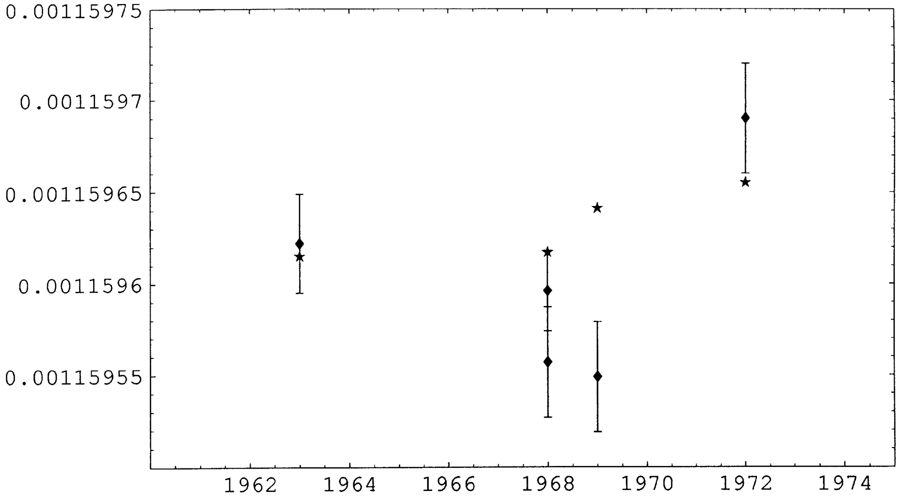

Fig. 5. *Plot of the 1963 Wilkinson and Crane result for the electron g-factor anomaly and subsequent re-evaluations. Theoretical values cited in these papers are also plotted. These values shifted during this period partly due to changes in the fine structure constant and partly due to refinements in theory.*

The precision of the theoretical result was limited by the still unknown third radiative correction (amounting to 10 on the last digits) and the current experimental error in the fine structure constant (5 on the last digit) which goes into the calculation. In the end of their paper, Wilkinson and Crane state that 'partly for this reason, but mainly for experimental reasons, we here conclude the 10-year effort of the laboratory on the g-factor of the free negative electron'. Nevertheless, just seven years later a new experiment with a further order of magnitude improvement in precision was reported from the same laboratory (Wesley and Rich, 1970).

In the meantime re-evaluations appeared of the Wilkinson and Crane experiment worsening the agreement between theory and experiment. Farley (1968) pointed out a theoretical error in the relativistic calculation of electron motion, Rich (1968) improved the numerical precision in averaging the actual magnetic field, and Henry and Silver (1969) made further relativistic corrections. Finally, in 1972 Granger and Ford made a careful re-evaluation of the theoretical basis for spin motion in the experiment.[9] In 1971 a significant change also happened in the theoretical prediction of the free electron anomaly due to the first calculation of the third radiative correction by Levine and Wright (1971). As seen in Fig. 5 the first corrections tended to worsen the agreement between theory and experiment whereas the Granger and Ford re-evaluation comfortably agreed with the Levine and Wright result.

[9] The re-evaluation relies entirely on non-quantum relativistic theory.

In 1970 Wesley and Rich (1970, 1971) rebuilt the spin-precession experiment allowing for an order of magnitude increase of the magnetic field strength along with other improvements. The increased magnetic field diminished the relative size of the correction due to stray electric fields. The stronger magnetic field also allowed the electrons to be trapped for millions of orbit revolutions leading to a smaller error on the $g-2$ frequency ω_a. The final result $a = 0.0011596577(35)$ agreed perfectly with the current theoretical value $a = 0.001159655(2)$ by Levine and Wright (1971). It almost agreed with the original Wilkinson and Crane value of $a = 0.001159622(27)$ within the quoted errors, but disagreed significantly with all but the last of the later re-evaluations of this experiment. The authors expressed worries about the disagreement, but wrote that in spite of an extensive critical review 'no concrete basis for the discrepancy has yet been found'.

In fact, Granger and Ford (1972) were able to explain also this discrepancy in their re-evaluation. Nevertheless when Rich and Wesley (1972) reviewed the situation later that year, they wrote (1972, p. 255):

> The agreement [between theory and experiment] should be treated with a certain amount of caution, since it is based on a comparison between a single theoretical calculation and a single type of experimental measurement. In view of the complexities of the theoretical calculation, and the difficulty of accurately estimating the systematic errors associated with a specific experiment, independent checks of both theory and experiment are of great importance.

At the end of the period of free electron precession experiments there was essentially only one experiment (Wesley and Rich, 1971) with several interpretations (Rich, 1968; Farley, 1968; Henry and Silver, 1969; Granger and Ford, 1972), and one theoretical calculation (Levine and Wright, 1971) at the highest level of precision. Apparently, this situation was considered uncomfortable.

7. Free Electron Spin Resonance Experiments

The third kind of experiments has its origin in an early experiment by Dehmelt (1958). Although the experimental precision was too small to compete with the atomic level experiments of that time, his method also avoided the binding corrections that limited the atomic experiments.

Free electrons in a magnetic field have on top of the energy levels associated with orbital motion two distinct levels, corresponding to spin up and down. The level separation between these spin states is given by $\hbar\omega_s$ where ω_s is the spin-flip frequency $\omega_s = geB/2m$, which is proportional to both the g-factor of the free electron and the magnetic field B. In a magnetic field free electrons will tend to become polarised by aligning their spin directions with the field in a relaxation time depending on the environment of the electrons. Subjecting the aligned electrons to a magnetic field oscillating at a frequency in the neighbourhood of the spin-flip frequency, the electrons become depolarised, and the strongest depolarisation happens exactly at the spin-flip frequency, a phenomenon akin to acoustic resonance.

In 1968, Gräff, Major, Roeder and Werth continued the work started by Dehmelt on free electrons. The experimental resolution could not compete with the precision of the precession experiments by Wilkinson and Crane (1963), but the experiment demonstrated (Gräff, Klempt and Werth, 1969) the feasibility of a direct resonance measurement of the anomaly. In the following we shall describe the principle of this experiment that forms the basis for all subsequent spin resonance experiments.

As already mentioned, a non-relativistic electron in a homogenous magnetic field B will move in a circular orbit with cyclotron frequency $\omega_c = \omega_0 = eB/m$. The radius of the orbit is given by $r = v/\omega_c$ and is proportional to the electron's velocity v (and inversely proportional to the magnetic field). This means that if the velocity is lowered, the radius of the orbit is shrunk correspondingly. In classical (non-quantum) mechanics there is no lower limit to this phenomenon, and an electron in a circular orbit will in analogy with a classical atom lose energy to electromagnetic radiation and spiral in towards the center. Classically, the electron velocity will grow towards infinity in the atom whereas in the magnetic field it will decrease towards zero.

Quantum mechanically both cases are, however, impossible because of the uncertainty relations which state that one cannot at the same time determine the velocity and the position of a particle with arbitrary precision. Consequently, for an electron in a magnetic field, like for an electron in an atom, there will be a lowest level of energy, a ground state, below which the electron cannot be found. Above the ground state there will be an infinite sequence of states, which for the electron in the homogenous magnetic field forms a ladder of equidistant levels, called Landau levels.[10]

The distance between the Landau levels is (in the non-relativistic case) given by the cyclotron frequency. If circulating electrons are subjected to oscillating electromagnetic fields of this frequency, transitions to higher Landau levels with larger radius will occur. Eventually the electrons may collide with the surrounding material. In their first experiment Gräff *et al.* (1968) used this effect to determine the cyclotron frequency.

The electron's spin only slightly changes this simple picture in a way which is reminiscent of the precession case, although the physics is quite different.[11] If the g-factor of the electron were equal to 2, the spin-flip frequency $\omega_s = geB/2m$ would be equal to the cyclotron frequency ω_c, and an electron with spin up in

[10] Because of the similarity between the electron's behaviour in an atom and in a magnetic field and since the field in these experiments may be viewed as 'anchored to the earth', the magnetically bound electron has been called geonium (Van Dyck, Schwinberg and Dehmelt, 1977).

[11] In the precession experiments the electron spin direction rotates at almost the same rate as the velocity direction, whereas in the resonance experiments the spin-flip level spacing is almost equal to the Landau level spacing. Loosely speaking, one may say that the electron precession experiments were classical, whereas the electron resonance experiments belong to the quantum realm.

a given Landau level would have precisely the same energy as an electron with spin down in the next higher Landau level. Due to the g-factor anomaly, this is not strictly the case and there is a small difference in energy between the two situations. The frequency corresponding to this difference is $\omega_a = \omega_s - \omega_c = a\omega_0$, which is directly proportional to the anomaly. The anomaly may thus be determined from the same formula as in the precession experiments:

$$a = \frac{\omega_a}{\omega_0}. \tag{5}$$

Technically, an important advance entered in this experiment. The electronic orbits may drift along the magnetic field lines. In order to contain the electrons, Gräff *et al.* employed a so-called Penning trap, in which an electric field is superimposed on the magnetic field. The electric field is generated by two negatively charged electrodes repelling the negatively charged electrons from the end of the containment region, together with a positively charged cylindrical electrode surrounding it. Although attracted by the positive cylinder, the electrons are prevented from moving towards it by the circular motion imposed by the magnetic field, as long as the electrode voltage is not too high. The levels are influenced by the imposed voltage, and in the end Gräff *et al.* (1969) could extrapolate to zero voltage in order to extract the desired frequencies.

In 1973 Walls and Stein employed a different technique for monitoring the two frequencies in the Penning trap. The slow axial oscillation of the electron orbits along the direction of the magnetic field gives rise to a noise in the circuit connected to the end plate electrodes. The amplitude of this noise is coupled to the polarisation of the trapped electrons, and by monitoring the noise around the spin-flip resonance and the anomaly resonance, the frequencies may be extracted. As in the preceding experiments the value for the anomaly was not competitive with the electron precession experiments of the time.[12] Problems with understanding the line widths also played a major role in this experiment, but the influence of the electric fields from the spatially distributed cloud of trapped electrons seemed to ultimately limit experiments of this particular type (see Van Dyck, 1990, p. 326).

The space charge problem would however be absent, if it were possible to trap a single electron and keep it under controlled conditions. In 1973 this was achieved in an extraordinary experiment (Wineland *et al.*, 1973) based on the observation that electrons in a Penning trap effectively behave like an electronic circuit resonating at the natural axial oscillation frequency of the trapped

[12] Walls reported already in 1970 a preliminary value using this technique for the anomaly in his Ph.D. thesis. As the result $a = 0.001159580(80)$ was never published, we have omitted it in Table 1 (see Rich and Wesley, 1972). It agrees but does not compete with previous experiments and theory at the time.

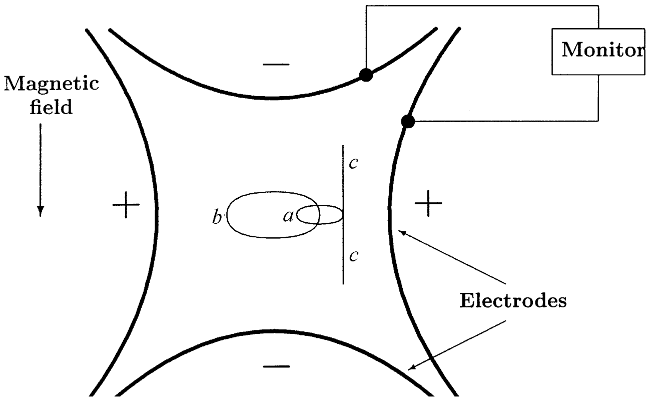

Fig. 6. *Sketch of (late) free electron spin resonance experiment for determination of the electron's g − 2. The Penning trap is a cavity with a homogeneous magnetic field superposed on a quadrupole electric field, defined by electrodes shaped as a hyperboloid of revolution (with main axis parallel to the magnetic field). In classical language, the electron performs a complicated motion which is a combination of a fast circular cyclotron motion around the field lines (a), a slow drift of the center of the cyclotron orbits around the axis of the cavity (b), and a slow longitudinal oscillation up and down the magnetic field lines (c). Furthermore, by means of a weak inhomogeneous magnetic field (not shown here), the frequency of the longitudinal oscillations can be made to depend on the electron's state. The longitudinal oscillation frequency, and therefore the electron's state, is monitored by an electronic resonance circuit.*

electrons (see Fig. 6). When brought into forced oscillations by means of a driving potential at a frequency near the natural one, the amplitude of the current in the circuit depends on the difference of the applied driving frequency and the resonance frequency, so that the response current is strongest closest to resonance.

The response is also proportional to the number of electrons in the trap. In the experiment it was demonstrated that it was possible to trap a few electrons and, by observing the circuit current, follow how they one by one were ejected from the trap when driven by a sufficiently strong potential. The last electron was held for a few minutes but that time could be made much longer by lowering the driving potential.

In 1977 this technique lead to a new high precision value (Van Dyck, Schwinberg and Dehmelt, 1977) for the electron anomaly with a quoted uncertainty seventeen times smaller than the best spin precession result. By perturbing the large homogenous magnetic field with a small bottle shaped magnetic field, the resonance frequency in the axial motion could be made dependent on the quantum state of a single trapped electron. The changes were minute but could be observed through the changes in the circuit response. This made possible the essentially simultaneous determination of both the cyclotron frequency ω_c and the anomaly frequency ω_a.

At that time there were three different theoretical calculations of the third order term in the anomaly, leading to three different predictions for the theoretical value. The experimental value fell almost within the ranges of these predictions, which had a spread over about five times the experimental error (Van Dyck, Schwinberg and Dehmelt, 1977).

Over the next decade Van Dyck, Schwinberg and Dehmelt (1979, 1984, 1987) refined this method to yield a further reduction of the uncertainty by a factor of fifty, yielding a final value of $a = 0.001159652189(4)$. The quoted error in the anomaly is now four parts per billion with the statistical error being much smaller than the systematic error. The systematic error is dominated by the influence of the cavity walls on the resonances, effects that have so far only been estimated.

The smallness of the statistical errors make the determination of cavity shifts paramount for getting even better experimental values for the electron anomaly. The authors themselves worry (Van Dyck, 1990) whether their previous work could be 'plagued by large cavity shifts'. As of 1997 no new experimental results have been reported, but new experiments are underway.[13]

It should be noted that this experiment is carried out by a single group in the world and thus lacks the dialogue with competing experiments. In this sense the experimental situation appears not much different from the one at the end of the period of the precession experiments.

Theoretically the calculation of the electron anomaly has been carried through to fourth order in the fine structure constant (Kinoshita, 1995). The first three orders are now known analytically, whereas the fourth has been evaluated numerically including muonic, weak and strong contributions. The intrinsic error in the theoretical calculation is about four times smaller than the current experimental error. In calculating a theoretical value it is however necessary to employ a value for the fine structure constant, but the problem is that the error in this quantity carries through to the error in the anomaly which thereby becomes several times larger than the experimental error. Thus, in the current situation a limit has been reached in the comparison between theory and experiment.[14]

8. Discussion

In the above account we have highlighted some of the places where inspiration from the theory under scrutiny, QED, *might* have been an important factor for the experimentalists when they established their final results. We emphasise

[13] Van Dyck, private communication.

[14] Furthermore there are currently three different determinations of the fine structure constant which fall outside each other's error bars (Kinoshita, 1996). There are thus three different theoretical predictions of the electron anomaly, all of which disagree with the experimental value, although the disagreements are all only a few times the quoted errors.

again the uncertainty on this point since, as already indicated, the published papers cannot be taken to represent the full range of the experimentalists' motivations. Nevertheless, the published results indicate what the authors can be expected to defend. Moreover, by seeing these experiments in a historical perspective, it is possible to extract some important lessons on the interplay between theory and experiment.

Consider the question of the theory-ladenness of the experiments. In the atomic level experiments non-relativistic quantum mechanics was the basic theory, but QED actually influenced the extraction of a $g-2$ value from the atomic experiments. Kusch and Foley (1947) needed the input from Schwinger to exclude the possibility that there could be an anomalous orbital g-factor. Thus theory in this case evidently directed the interpretation of the experiment in clearing up an ambiguity, but otherwise QED was not needed. As we indicated earlier, relativistic (non-QED) mass corrections were also applied to later experiments, in order to obtain a value for the g-factor of the free electron, which could be compared with theory.

In the spin precession experiments, classical relativistic treatment was sufficient to calculate the precession rate, whereas quantum mechanics only came marginally in through preparation and analysis of the electrons by means of Mott scattering. Spin resonance experiments again relied entirely on non-relativistic quantum mechanics. Thus, in these experiments, QED was not involved in the extraction of a g-value from experiment.

But a theory under test can also in other ways influence an experiment. In connection with the ending of the gyromagnetic experiments, Galison provides the following conjecture (1987, p. 74):

> One might expect that in experiments where both strong theoretical predispositions and a definite quantitative prediction are present, it will often happen that the experimenter will end an experiment by finding the anticipated result, whether or not it corresponds with what is later found to be the case.

It is clear that the $g-2$ experiments (except for the very first) were pursued in the specific theoretical environment of QED. Without this theory there would not have been much point in pushing the experiments to higher and higher precision. But to what extent can history support a hypothesis that the theoretical environment of QED prompted the experimenters to get agreement with theory (e.g. by looking for systematic errors until an agreement was obtained)? Let us summarise the three phases of the $g-2$ experiments with this question in mind.

For the first atomic resonance experiment by Kusch and Foley there was only Breit's theoretical estimate for the g-factor of the electron, and Kusch and Foley's result was in disagreement with it. The following atomic resonance experiments initially agreed with theoretical calculations (by Schwinger, and Karplus and Kroll), but the last in this series (by Franken and Liebes) did not conform with theory. In turn, this led to re-examinations of the theoretical prediction revealing an error in the previous calculation. Accordingly, theoretical bias can hardly be blamed as the main factor in reporting the final result. While the anomalous magnetic moment of the electron 'is one of the simplest quantities precisely calculable from first principles' (Kinoshita, 1995), the history of the calculation shows that this does not imply simplicity in determining its actual value.[15]

The free electron spin precession experiments with highest accuracy agreed in all cases with the theoretical prediction when they were published. However, the experimenters were not without reservations concerning this agreement. Schupp, Pidd and Crane (1961) worried about the reasonableness of their systematic corrections. Though Wilkinson and Crane were content enough with the situation so as to consider it to be the end of this series of the experiments, their laboratory was back in the same business only seven years later. At this time it was Wesley and Rich (1970, 1971) who came up with a more precise value and at the same time expressed clear concerns that their value did not agree with that of Wilkinson and Crane (or its three re-evaluations Farley (1968), Rich (1968), Henry and Silver (1969)). Even when the struggle over the systematic errors in Wilkinson and Crane's experiment had ended with Granger and Ford's (1972) reanalysis, Rich and Wesley were uneasy with the situation, because there was only one experimental value at the highest precision, only one equally precise theoretical calculation, and only one analysis of the systematic errors which brought theory and experiment into complete agreement (Rich and Wesley, 1972).

Thus, even if knowledge of the theoretical result may have played a role in these experiments, the experimenters were aware of the problem, and did in fact worry about too good an agreement (Rich and Wesley, 1972) based on a single theoretical paper (Granger and Ford, 1972). Moreover, as we saw in Fig. 5 the first three re-evaluations of Wilkinson and Crane's result tended to shift the experimental $g-2$ value away from the theoretical value.

The free electron spin resonance experiments only became competitive with the best free electron spin precession experiments after 19 years (recall Table 1). Since that time (1977) the Van Dyck-Schwinberg-Dehmelt group has been the only one reporting new measurements of $g-2$. Without exceptions these results have been in agreement with theory although, as we mentioned above, the current situation does not really permit detailed comparison between theory and experiment. Nevertheless, the Van Dyck-Schwinberg-Dehmelt group continues to work on getting the systematic errors under better control.

We now turn to the question of trust in experimental results in the light of the above discussion. An often cited criterion for belief in experimental results is its stability under variation of the experimental circumstances (Radder, 1996). By

[15] The calculation of higher order terms in the fine structure constant becomes increasingly complicated in higher orders. The numbers of Feynman diagrams involved in the calculation from first to fourth order in α are 1, 7, 72, and 891. The first three orders have been evaluated analytically, whereas the fourth has only been calculated numerically (Kinoshita, 1995).

this criterion the result of Kusch and Foley from 1947 has been amply confirmed by a number of different experimental methods. By the same criterion we can have less confidence in the last three digits of the present $g-2$ value than in the first few digits.

Besides trust, the historical development of the experiments also supports a certain kind of realism. In *Representing and Intervening*, Hacking (1983) argues that we ought to believe in the existence of electrons, as these can be manipulated to study other phenomena. Hacking distinguishes between experimentation *with* electrons from experimentation *on* electrons, so that the former but not the latter constitute belief in the existence of electrons (Hacking, 1983, p. 265). In the case of $g-2$ experiments, Hacking's criterion of manipulability seems especially relevant for Dehmelt and coworkers' deft handling of single electrons, while the distinction between experimenting with and experimenting on is harder to maintain. Hacking's distinction may be saved, however, if one allows for the possibility of experimenting with well-known properties of electrons to study other less well-known properties of electrons.

Seen as a whole, the historical development consists in a gradual stripping away of the electron's environment, with a corresponding elimination of systematic errors. In the atomic resonance experiments the electrons were found deep inside the atoms, making the extracted value dependent on complicated atomic-physics calculations. In the free electron spin precession experiments the electrons were removed from the atom and studied collectively in a magnetic field trap, but space charge problems due to the cloud of electrons ultimately set the limit to this kind of experiments. Finally, in the single electron spin resonance experiments the electrons in a Penning trap could eventually be controlled so well as to eject all but one of them from the trap.

9. Conclusion

In our view the historical progression of the experiments not only speaks in favour of the trust in the experimental results, but also supports the existence of electrons through their sublime manipulability. Thus, insofar as there are electrons, they have an anomalous magnetic moment.

We have not proven that the data analysis in the $g-2$ experiments was not influenced by knowledge of the QED predictions. However, we find it implausible that this should be the case due to the long sequence of $g-2$ experiments with their continuing stripping of the electron's environment. This stripping process was entirely based on theory that did not involve QED itself. QED was evidently a driving force for the historical development of the $g-2$ experiments, but apart from our reservations concerning the atomic level experiments, it played essentially no role in the design and execution of the experiments.

The trust in the results constitutes a clear empirical success for QED. Whether this implies that QED is necessarily the correct framework for describing the electron is another story.[16] In any case, a different theory would have to face up to the remarkable results for the anomalous magnetic moment of the electron.

References

Beringer, R. and Heald, M. (1954) 'Electron Spin Magnetic Moment in Atomic Hydrogen', *Physical Review* **95,** 1474–1481.
Breit, G. (1947) 'Does the Electron Have an Intrinsic Magnetic Moment?', *Physical Review* **72,** 984.
Collins, H. M. (1984) '*When* do Scientists Prefer to Vary their Experiments?', *Studies in History and Philosophy of Science* **15**, 169–174.
Combley, F. H. (1979) '$(g-2)$ Factors for Muon and Electron and the Consequences for QED', *Reports on Progress in Physics* **42,** 1889–1935.
Dehmelt, H. G. (1958) 'Spin Resonance of Free Electrons Polarized by Exchange Collisions', *Physical Review* **109,** 381–385.
Dirac, P. A. M. (1928) 'The Quantum Theory of the Electron', *Proceedings of the Royal Society* **A117,** 610–624.
Dyck, R. S. van, Schwinberg, P. B. and Dehmelt, H. G. (1977) 'Precise Measurements of Axial, Magnetron, Cyclotron, and Spin-Cyclotron Beat Frequencies', *Physical Review Letters* **38,** 310–314.
Dyck, R. S. van, Schwinberg, P. B. and Dehmelt, H. G. (1979) 'Progress of the Electron Spin Anomaly Experiment', *Bulletin American Physics* **24,** 758.
Dyck, R. S. van, Schwinberg, P. B. and Dehmelt, H. G. (1984) 'The Electron and Positron Geonium Experiments', in R. S. van Dyck and E. N. Fortson (eds), *Atomic Physics 9* (Singapore: World Scientific), pp. 53–73.
Dyck, R. S. van, Schwinberg, P. B. and Dehmelt, H. G. (1987) 'New High-Precision Comparison of Electron and Positron g Factors', *Physical Review Letters* **59,** 26–29.
Dyck, R. S. van (1990) 'Anomalous Magnetic Moment of Single Electrons and Positrons: Experiment', in T. Kinoshita (ed.), *Quantum Electrodynamics* (Singapore: World Scientific), pp. 322–388.
Farago, P. S., Gardiner, R. B., Muir, J. and Rae, A. G. A. (1963) 'Direct Measurement of the g-Factor Anomaly of Free Electrons', *Proceedings of the Physical Society of London* **82,** 493–499.
Farley, F. J. M. (1968) '$(g-2)$ Experiments on the Electron and Muon and Related Topics', in M. Lévy (ed.), *Cargèse Lectures in Physics, Vol. 2* (New York: Gordon and Breach), pp. 55–117.
Foley, H. M. and Kusch, P. (1948) 'On the Intrinsic Moment of the Electron', *Physical Review* **73,** 412–413.
Franken, P. and Liebes, S. (1956) 'Magnetic Moment of the Proton in Bohr Magnetons', *Physical Review* **104,** 1197–1198.
Franklin, A. (1986) *The Neglect of Experiment* (Cambridge: Cambridge University Press).
Galison, P. (1987) *How Experiments End* (Chicago: Chicago University Press).
Gardner, J. H. and Purcell, E. M. (1949) 'A Precise Determination of the Proton Magnetic Moment in Bohr Magnetons', *Physical Review* **76,** 1262–1263.

[16] Schwinger, who initiated the theoretical calculations of $g-2$ from QED, later became very sceptical towards aspects of the QED framework, leading him to his alternative *source theory* (Schwinger, 1970). Some discussion of source theory which differs from QED with respect to the physical interpretation of the internal radiation processes associated with renormalisation may be found in (Rugh, Zinkernagel and Cao, 1999).

Gräff, G., Major, F. G., Roeder, R. W. H. and Werth, G. (1968) 'Method for Measuring the Cyclotron and Spin Resonance of Free Electrons', *Physical Review Letters* **21,** 340–342.

Gräff, G., Klempt, E. and Werth, G. (1969) 'Method for Measuring the Anomalous Magnetic Moment of Free Electrons', *Zeitschrift für Physik* **222,** 201–207.

Granger, S. and Ford, G. W. (1972) 'Electron Spin Motion in a Magnetic Mirror Trap', *Physical Review Letters* **28,** 1479–1482.

Hacking, I. (1983) *Representing and Intervening* (Cambridge: Cambridge University Press).

Henry, G. R. and Silver, J. E. (1969) 'Spin and Orbital Motions of a Particle in a Homogenous Magnetic Field', *Physical Review* **180,** 1262–1263.

Karplus, R. and Kroll, N. M. (1950) 'Fourth-Order Corrections in Quantum Electrodynamics and the Magnetic Moment of the Electron', *Physical Review* **77,** 536–549.

Klapdor-Kleingrothaus, H. V. and Staudt, A. (1995) *Non-accelerator Particle Physics* (London: Institute of Physics Publishing).

Kinoshita, T. (1995) 'New Value of the α^3 Electron Anomalous Magnetic Moment', *Physical Review Letters* **75,** 4728–4731.

Kinoshita, T. (1996) 'The Fine Structure Constant', *Reports on Progress in Physics* **59,** 1459–1492.

Koenig, S. H., Prodell, A. G. and Kusch, P. (1952) 'The Anomalous Magnetic Moment of the Electron', *Physical Review* **88,** 191–199.

Kusch, P. and Foley, H. M. (1947) 'Precision Measurements of the Ratio of the Atomic g Values in the $^2P_{3/2}$ and $^2P_{1/2}$ States of Gallium', *Physical Review* **72,** 1256–1257.

Kusch, P. and Foley, H. M. (1948) 'The Magnetic Moment of the Electron', *Physical Review* **74,** 250–263.

Lautrup, B., Peterman, A. and Rafael, E. de (1972) 'Recent Developments in the Comparison Between Theory and Experiments in Quantum Electrodynamics', *Physics Reports* **3C,** 196–253.

Levine, M. J. and Wright, J. (1971) 'Sixth-Order Magnetic Moment of the Electron', *Physical Review Letters* **26,** 1351–1353.

Louisell, W. H., Pidd, R. W. and Crane, H. R. (1953) 'An Experimental Measurement of the Gyromagnetic Ratio for the Free Electron', *Physical Review* **91,** 475.

Louisell, W. H., Pidd, R. W. and Crane, H. R. (1954) 'An Experimental Measurement of the Gyromagnetic Ratio for the Free Electron', *Physical Review* **94,** 7–16.

Mott, N. F. (1929) 'The Scattering of Fast Electrons by Atomic Nuclei', *Proceedings of the Royal Society of London* **A124,** 425–442.

Nafe, J. E., Nelson, E. B. and Rabi, I. I. (1947) 'The Hyperfine Structure of Atomic Hydrogen and Deuterium', *Physical Review* **71,** 914–915.

Peterman, A. (1957) 'Magnetic Moment of the Electron', *Nuclear Physics* **3,** 689–690.

Particle Data Group, (1994) 'Review of Particle Properties', *Physical Review* **D 50,** 1173–1825.

Radder, H. (1996) *In and About the World — Philosophical Studies of Science and Technology*, (New York: New York State University Press).

Rich, A. (1968) 'Corrections to the Experimental Value for the Electron g-Factor Anomaly', *Physical Review Letters* **20,** 967–971. Erratum: *Physical Review Letters* **20,** 1221–1222 (1968).

Rich, A. and Wesley, J. C. (1972) 'The Current Status of the Lepton g Factors', *Reviews of Modern Physics* **44,** 250–283.

Rugh, S. E., Zinkernagel, H. and Cao, T. Y. (1999) 'The Casimir Effect and the Interpretation of the Vacuum', *Studies in History and Philosophy of Modern Physics* **30,** 111–139.

Schupp, A. A., Pidd, R. W. and Crane, H. R. (1961) 'Measurement of the g Factor of Free, High-Energy Electrons', *Physical Review* **121,** 1–17.

Schweber, S. S. (1994) *QED and the men who made it: Dyson, Feynman, Schwinger, and Tomonaga* (Princeton: Princeton University Press).

Schwinger, J. (1948) 'On Quantum-Electrodynamics and the Magnetic Moment of the Electron', *Physics Review* **73,** 416–417.

Schwinger, J. (1970) *Particles, Sources and Fields, Vol. 1* (Reading, Mass.: Addison–Wesley).

Sommerfield, C. M. (1957) 'Magnetic Dipole Moment of the Electron', *Physics Review* **107,** 328–329.

Uhlenbeck, G. E. and Goudsmit, S. (1926) 'Spinning Electrons and the Structure of Spectra', *Nature* **117,** 264–265.

Walls, F. L. and Stein, T. S. (1973) 'Observation of the $g-2$ Resonance of a Stored Electron Gas Using a Bolometric Technique', *Physical Review Letters* **31,** 975–979.

Wesley, J. C. and Rich, A. (1970) 'Preliminary Results of a New Electron $g-2$ Measurement', *Physical Review Letters* **24,** 1320–1325.

Wesley, J. C. and Rich, A. (1971) 'High-Field Electron $g-2$ Measurement', *Physical Review* **A4,** 1341–1363.

Wilkinson, D. T. (1962) Ph.D. Thesis, *University of Michigan, University Microfilms* 63–478.

Wilkinson, D. T. and Crane, H. R. (1963) 'Precision Measurement of the g Factor of the Free Electron', *Physical Review* **130,** 852–863.

Wineland, D., Ekstrom, P. and Dehmelt, H. G. (1973) 'Monoelectron Oscillator', *Physical Review Letters* **31,** 1279–1282.