Some thoughts on shielding, penetration and effective nuclear charge.

John E. McGrady*

Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QR, U.K.

E-mail: john.mcgrady@chem.ox.ac.uk

Abstract

The concept of an effective nuclear charge (Z^{eff}) is one of the foundations of inorganic chemistry, yet at the same time a source of great confusion. A case in point is the $4s/3d$ orbital ordering in the first transition series: the bare facts are not in dispute - K has a $4s^1$ ground-state configuration while isoelectronic Ti^{3+} is $3d^1$ - but the precise reasons why this change in ground state configuration occurs have been hotly debated. The Schrödinger equation tells us that the atomic or ionic ground state is, ultimately, governed by a balance between kinetic energy, Coulomb attractions between electrons and the nucleus and Coulomb repulsions between the electrons. In defining an effective nuclear charge, however, we conceal the mathematically challenging electron-electron repulsions by offsetting them against some of the attractive electron-nucleus interaction. This ansatz then promotes the erroneous view that fluctuations in Z^{eff} are really due to differences in the electron-nucleus interaction, simply because this is the only term that survives in the model Hamiltonian. It is then a short leap to conclude that the way to maximise Z^{eff}, and hence the stability of the electron, is for it to approach the nucleus as closely as possible to maximise the Coulomb attraction. However, this ignores the fact that the core electrons are also close to the nucleus, and so

the electron-electron repulsions will be minimised when the electron in question is furthest from the nucleus, not when it is closest! The charge on the nucleus is a constant, so variations in Z^{eff} are simply proxies for variations in electron-electron repulsions: higher Z^{eff} really means less electron-electron repulsion. It is, therefore, variations in electron-electron repulsion, and not variations in the electron-nucleus attraction, that we must consider when we seek physical meaning in the effective nuclear charge. In teaching this central topic, we should place a greater emphasis on what is real (electron-electron repulsion) and not what is mathematical convenience (an artificially reduced nuclear charge). Only in this way can the real physical origins of trends in orbital stability be exposed.

Introduction

The idea that an electron experiences an "effective nuclear charge", somewhat reduced from the true nuclear charge as a result of shielding by the other electrons, can be traced back over a century to Moseley's work on the X-ray spectra of atoms. These spectra typically feature 2 series of lines, the K and L series, which are now understood to correspond to transitions from outer electron shells to vacancies in core states with principal quantum numbers $n = 1$ and $n = 2$. Moseley established that the frequencies, ν, of both series could be expressed as:

$$\nu = C_n \left(Z - \sigma\right)^2 \tag{1}$$

where Z is the atomic number and $\sigma = 1.0$ and 7.4 for the K and L lines, respectively. Moseley used this relationship to establish that it is the atomic number, rather than the atomic weight, which determines the position of a given element in the periodic system. Whilst the values of σ emerged, in the first instance, simply as empirical fitting parameters, it is now clear that the higher value of 7.4 for the L series reflected the shielding of the coulomb potential by the inner $1s^2(2s,2p)^6$ core electrons. Soon after the development of Schrödinger's wave theory a decade later, Zener optimised the energies of the first row atoms

Li-Ne by treating the exponent ζ in radial wavefunctions of the form $r^{n-1}e^{-\zeta r}$ as a variational parameter. The average orbital energy for the atoms of the first period is given (in units of eV) by:

$$E = -13.6\left(\frac{Z^{eff}}{n}\right)^2$$

where Z^{eff}, the 'mean effective charge', varies from 1.26 for Li to 5.88 for Ne. Slater took a rather simpler approach to the same problem, adjusting the values of the shielding constants, σ, in radial wavefunctions of the form:

$$R(r) = r^{(n-1)}e^{-(Z-\sigma)/nr} \tag{2}$$

to match observed spectroscopic parameters. Slater's study encompassed the majority of the periodic table and, critically, allowed him to develop a set of global empirical rules for computing the shielding constants for a given electron, i, in any atom. The effective nuclear charge for the ith electron in an atom is then calculated using Equation 3, where the summation runs over the j other electrons. The structure of the equation therefore reflects the important physical point that an electron does not shield itself. The subscript i in Equation 3 is an important reminder that the effective nuclear charge is a property of an individual electron in an atom, not a property of the atom or ion as a whole.

$$Z_i^{eff} = Z - \sigma = Z - \sum_{j \neq i} \sigma_{ij} \tag{3}$$

Slater then stated a series of empirical rules that determine the exact values of the shielding constants. The algorithm goes as follows:

Divide the electrons up into sub-shells:

$$1s\ \{2s,2p\}\ \{3s,3p\}\ \{3d\}\ \{4s,4p\}\ \{4d\}\ \{4f\}\ \{5s,5p\}$$

1. If the reference electron is in an s or p orbital, the shielding constants are calculated as follows:
    (a) For an electron in the same sub-shell, σ_{ij}=0.35, except for the 1s sub-shell where it is 0.30
    (b) For an electron in a sub-shell with principal quantum number one less than the reference electron, σ_{ij}=0.85.
    (c) For an electron in a sub-shell with principal quantum number two or more less than the reference electron, σ_{ij}=1.0.
2. If the reference electron is in a d or f orbital, the shielding constants are calculated as follows:
    (a) For an electron in the same sub-shell, σ_{ij}=0.35.
    (b) For electrons in all sub-shells (either with lower values of n or of l), σ_{ij}=1.0.

I will comment later on the motivation for the numbers Slater used, but for now let us accept them as empirical parameters and see how they are used. As an example, let us calculate Z^{eff} for one of the 2p electrons of Ne (it does not matter which one!). Slater's rules assign shielding constants, σ_{ij}, of 0.35 for each electron in the same 2s,2p shell (7 in total, 2 in 2s and 5 in 2p) and 0.85 for the two 1s electrons, giving:

$$Z^{eff}(2p) = 10 - 7 \times 0.35 - 2 \times 0.85 = 5.85$$

compared to Zener's value of 5.88. According to Slater's protocol, the inner 1s electrons, in contrast, are not shielded by the outer 2s,2p shell at all, leading to a much larger Z^{eff}:

$$Z^{eff}(1s) = 10 - 8 \times 0.0 - 1 \times 0.30 = 9.70$$

Slater's rules are necessarily a broad-brush approximation – they do not differentiate $2s$ from $2p$, for example, nor do they account for shielding by outer electrons. More complex rules for computing shielding constants that can capture these more subtle effects have subsequently been published by Clementi and others, but the greater accuracy comes at the cost of greater complexity. The power of Slater's rules lie in their simplicity – they rely on only a very few parameters that can be remembered easily, so they provide the basis for many of the 'rules of thumb' that guide inorganic chemists. As a result, Slater's name remains synonymous with the concepts of shielding, penetration and effective nuclear charge almost a century after his original work. It is useful at this juncture to pause and consider why it is that, even in the modern era where *ab initio* and density functional calculations are readily accessible and, in many cases, routine for even quite large systems, simple concepts like effective nuclear charge remain indispensable in inorganic chemistry. Despite all the advances in electronic structure theory, the exact ground-state wavefunction for a poly-electronic atom remains an inaccessible entity. The Hamiltonian operator can certainly be formulated exactly in terms of its fundamental components, kinetic energy ($\hat{T}$), nucleus-electron attraction ($\hat{V}_{ne}$) and electron-electron repulsion ($\hat{V}_{ee}$), Equation 3, but analytical solutions are unavailable. In purely mathematical terms, the electron-electron repulsion component of the Hamiltonian is a 2-electron operator, and this prevents the separation of variables into a sum of 1-electron components. If this component is absent, as it is in 1-electron atoms/ions such as hydrogen, the equation is separable and then exact analytical solutions are accessible.

$$\hat{H} = -\tfrac{1}{2}\textstyle\sum_i \nabla_i^2 - \sum_i \frac{Z}{r_i} + \sum_i \sum_{j\neq i} \frac{1}{r_{ij}} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee} \tag{4}$$

The recognition that it is the electron-electron repulsion component of the Hamiltonian that causes all the problems immediately suggests a work-around: we can simplify the Hamiltonian dramatically by removing the mathematically difficult electron-electron repulsion term and offsetting its effects against a portion of the attractive electron-nucleus interaction. The phrase “offsetting” is used quite deliberately rather than the more common term “shielding” to highlight the critical point that the effective nuclear charge, whilst mathematically convenient, is nothing more than a proxy for the physically real phenomenon, which is the electron-electron repulsion. The resultant hydrogen-like Hamiltonian:

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i \frac{Z_i^{eff}}{r_i} = \hat{T} + \hat{V} \tag{5}$$

is then separable, and leads trivially to hydrogenic solutions for each electron with eigenvalues $E_i = -13.6(Z_i^{eff}/n_i)^2$. The two Hamiltonians shown in Equations 3 and 4 represent opposite ends of the quantum mechanical spectrum: the first is exact but intractable while the second is highly simplified but trivially soluble. In moving to the simplified version, it is important to appreciate that we pay a heavy price: we completely lose sight of the electron-electron repulsions. In the following paragraphs, I address two specific questions that exemplify the use (and indeed misuse) of the effective nuclear charge concept in inorganic chemistry. The first of these is the textbook question of the relative stabilities of the $2s$ and $2p$ orbitals: why is the $2s$ orbital is more stable than $2p$ in all poly-electronic atoms? The question of why the ground-state configuration of Li is $2s^1 2p^0$ rather than $2p^1 2s^0$ will be familiar to every student of undergraduate-level chemistry. The second question, also much beloved of examiners, is why the ground state configuration of K is $4s^1 3d^0$ but for isoelectronic Ti^{3+} it is $3d^1 4s^0$. This observation lies behind the inorganic chemist’s rule of thumb that the $4s$ electrons are “lost first” when a transition metal is ionised, but textbook explanations are (in the authors opinion!) confusing. I approach these problems first using Z^{eff} as an empirical parameter, and then use Hartree-Fock theory to explore the physics that actually causes the differences. What really lies behind the relative stability of the $2s$ and $2p$ orbitals in the first row? Let us consider first the question of why the $2s$ orbital is more stable than $2p$ in the first row (and indeed any atom with more than one electron). The textbook answer will typically refer to the expression for the energy: $E = -13.6(Z^{eff}/n)^2$

and argue that Z^{eff} is greater for $2s$ than for $2p$, and so the electron is more stable in a $2s$ orbital. The more interesting question is to ask why Z^{eff} is greater for $2s$ than for $2p$. The radial distribution functions shown in Figure 1 appears to offer a facile explanation: the inner radial node of the $2s$ orbital lies very close to the nucleus – the radial maximum is in fact slightly closer than the radial maximum in the core 1s orbital. The argument then goes that the $2s$ orbital 'penetrates' inside the $1s^2$ core and therefore experiences a greater net attraction to nucleus than a $2p$ electron, which is largely outside the core. To quote from some popular textbooks: Purcell and Kotz "the $2s$ electron experiences a nuclear charge that is greater than Z-2" Shriver and Atkins (5th) "The closer to the nucleus an electron can approach, the closer is the value of Z^{eff} to Z because the electron is repelled less by other electrons present in the atom" … "There is a non-zero probability that a $2s$ electron can be found inside the 1s shell and experience the full nuclear charge" Housecroft and Constable "There are two types of electrostatic interaction at work".. (attraction and repulsion). "And electron in the $2s$ orbital will spend more time nearer the nucleus than an electron in the $2p$ orbital". Burrows, Holman et al "…there is an area of electron density relatively close to the nucleus for the $2s$ electron. The $2s$ electron is said to penetrate the 1s electrons and thus a $2s$ electron feels a higher effective nuclear charge than a $2p$ electron". Note how the emphasis is all on the very small amount of electron density that lies inside the radial node of the $2s$ orbital (the small blue peak in Figure 1). What is the justification for focussing on this rather than the very much large peak, which contains most of the electron density? An obvious objection to this argument is that the coulomb repulsion between two particles is not directional – it depends only on the distance between two particles, so just because the $2s$ electron spends time closer to the nucleus than the 'core' 1s electrons, does not mean that it is not repelled by them! Indeed, we could make the counter argument that the inner radial maximum of the $2s$ orbital is a very bad place for an electron to be, because it is very close to the core 1s electrons.

A review of some basic electrostatics highlights many of the important points in the

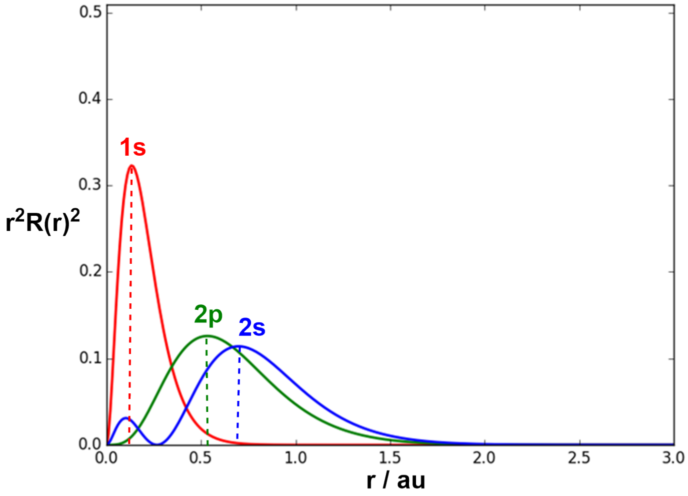

Figure 1: Radial distribution functions in a generic first-row atom.)

following discussion. Coulomb's law states that the electrostatic potential at a point is the sum of the potentials due to all the charged species:

$$\hat{V}_i = -\sum_j \frac{q_i q_j}{r_{ij}} \tag{6}$$

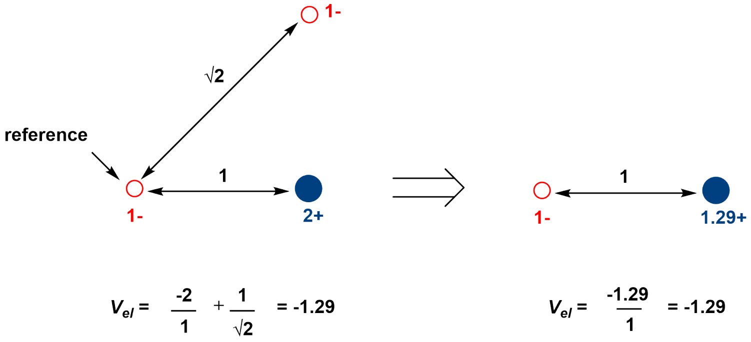

Figure 2: Electrostatic potentials of a 3-particle system and a model 2-particle system.)

An obvious corollary of this is that the introduction of a second electron does not change the interaction between a proton and electron (as long as the two original particles do not move). The total electrostatic energy certainly changes, but that is a different thing. So in the 3-particle system shown on the left in Figure 2 (2 electrons, one nucleus with a 2+ charge), the total potential energy of the reference electron is $V = -\frac{2}{1} + \frac{1}{\sqrt{2}} = -1.29$. Certainly, the net potential (–1.29) is less than it would be in the absence of the second electron (-2), but the interaction between the nucleus and the first electron is unchanged (at –2). In the spirit of the effective nuclear charge argument, we could create a fictitious one-electron model system with the same net potential (–1.29) by reducing the nuclear charge to +1.29. This seems, at first glance, to be a very promising strategy because we reproduce the observable property of the real system (the electrostatic potential) using a much simpler model. However, the pitfalls emerges as soon as we ask the simple question: where could we move the reference electron in order to make it more stable? If this question seems a little abstract, bear in mind that it is effectively the question we ask when we try to decide on the relative stability of $2s^1$ *vs* $2p^1$ in Li, or $4s^1$ *vs* $3d^1$ in K. In the model 1-electron system with Z = +1.29, there is only one solution – we have to move the electron closer to the nucleus. In the real system, in contrast, there are two solutions: we could move the electron closer to the nucleus, but we could also move it further away from the other electron. By concentrating our attention on the simpler model system (or, equivalently, on Z^{eff} in an atom), we lose sight of this second possibility, and it encourages us to imagine that close approach to the nucleus is the only way to increase stability. To bring the conversation back to atoms (and specifically poly-electronic atoms), we need to remember that stability is fundamentally a compromise between electrostatic attraction (to the nucleus) and electrostatic repulsion (with other electrons), as embodied in Equation 3. So whilst closer approach of an electron to the nucleus will certainly enhance the attraction, it does so only at the cost of greater repulsion with the core electrons. In fact, Slater's equation (Equation 2) captures this point beautifully – if we want to compare Z^{eff} for 2 electrons in the same atom, it is that varies, not Z!. It is reasonable to ask why the balance between attraction and repulsion is not usually emphasised in undergraduate textbook treatments of the subject (with the exception of Housecroft and Constable). The answer lies in the all-pervading influence of the concept of effective nuclear charge, which has dominated the thinking of inorganic chemists for generations. It has become so natural to explain periodic trends using Z^{eff} as the fundamental variable that the community has stopped thinking about what Z^{eff}actually is. If we look at the question of $2s^1$ *vs* $2p^1$ in Li, and we limit ourselves to thinking only in terms of maximising attraction, it is a very short leap to conclude that the inner radial maximum of $2s$ (Figure 1) must be the cause of its relative stability (as all of the major textbooks seem to do), simply because it affords the closest approach to the nucleus (the $2s$ electron 'penetrates' through the $1s^2$ shell more effectively than $2p$). If we instead consider the reality that higher Z^{eff} really means less electron-electron repulsion (Equation 2), and not more electron-nucleus attraction, we see that an electron in a $2s$ orbital can also gain stability by virtue of being further away from the core electrons in the $1s^2$ shell. How does this connect with the conventional ideas of penetration? The presence of the radial node in $2s$ certainly means that a small amount of the density lies close to the nucleus, but it also has another consequence: the outer (and much larger) radial maximum lies further away from the nucleus, and therefore also further away from the $1s^2$ core, than does the (only) radial maximum for the $2p$ orbital. The majority of the $2s$ density is therefore further from the $1s^2$ core, is repelled less by it, and this is the decisive factor in determining the order of filling ($2s$ before $2p$). In reality, the greater Z^{eff} for a $2s$ electron *vs* $2p$ at Li does not reflect more attraction of the $2s$ electron to the nucleus, it reflects less repulsion of the $2s$ electron by the $1s^2$ core. Note that the emphasis here is squarely on the outer radial maximum where the electron density is largest - this seems intrinsically more satisfactory than focussing on the very small amount of electron density close to the nucleus! Before moving on to the $4s/3d$ discussion, I acknowledge that I have provided no concrete evidence to support the argument that we should focus on repulsion as the central variable, and that the decisive feature of the $2s$ orbital is the position of the outer radial maximum in the radial distribution function, not the small maximum close to the nucleus (although I hope that the arguments are somewhat convincing!). It is possible to do so using Hartree-Fock theory, and I present these arguments in the attached appendix for the interested reader. The arguments are, however, somewhat complex and rely heavily on the deviation of orbitals from their strictly hydrogenic forms, which is never addressed in standard texts. They also require a reasonable grasp of quantum mechanical phenomena such as coulomb and exchange interactions. The appendix can safely be skipped at this stage by those less interested in the quantum mechanical fine details!

The $4s/3d$ ordering controversy in the fourth period

I now return to the rather controversial question of the $4s/3d$ orbital ordering, a topic which has provoked extensive debate in journals such as *J. Chem. Ed.* over the years. The bare facts here are that the ground state configuration of K (atomic number 19) is Ar$4s^1$ while for isoelectronic transition metals cations such as Ti^{3+}, it is Ar$3d^1$. In essentially parallel arguments to those made for Li, the preference for the $4s^1$ configuration at potassium is typically rationalised by the highly penetrating nature of $4s$, which allows it to 'experience more of the nuclear charge'. I will approach this problem first from the perspective of Z^{eff} and Slater's rules. I will then also use Hartree-Fock theory to attach some hard numbers to attraction and repulsion, to see what really lies behind the trends. I use Hartree-Fock theory here but not in the previous sections because the arguments are less complicated, and do not require an explicit consideration of the deviation from hydrogenic character of the orbitals.

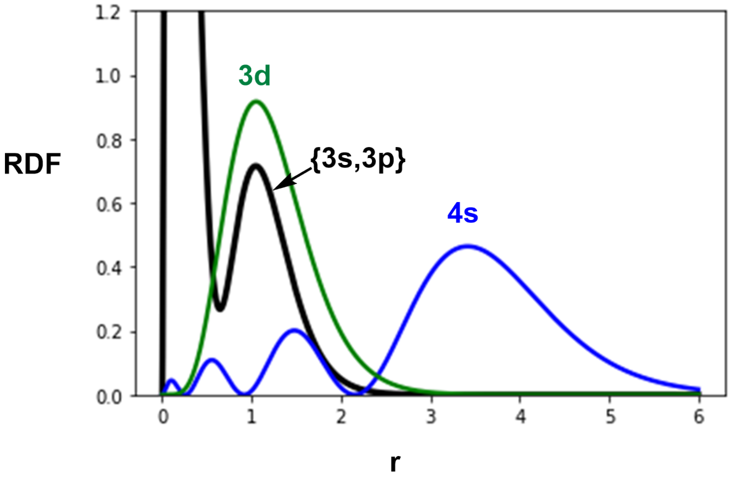

Figure 3: Radial distribution functions for the valence and core orbitals of a transition metal.)

Before starting, however, I wish to comment on the rather unhelpful statement often encountered in textbooks, that the $4s$ electrons are 'added first', in the sense that $4s$ is filled before $3d$ in the periodic table, but are also 'removed first', in the sense that when we ionise Ti ($4s^2 3d^2$) we get Ti^{3+} $3d^1$, not $4s^1$. This is unhelpful because 'first' means entirely different things in the two cases: in the periodic sense in which we add $4s$ 'first', we are starting from a neutral atom (Ar) and adding both a proton and an electron (and some neutrons) simultaneously to form K, Ca *etc.* When we ionise Ti, we remove an electron ($4s$ 'first') but leave the proton in place, and this is an entirely different process – there really is no reason why the two should be related! Look at the problem another way: we could form Ti^{3+} by adding 3 protons to a K nucleus – admittedly, we would need a particle accelerator to do so, but that does not alter the fact that the ground-state configuration will surely be the same as if we had produced Ti^{3+} conventionally by ionising atomic Ti. The ground state configuration of any atom/ion is a state function – it cannot possibly depend on how that atom/ion was formed, and so it makes little sense to attempt to rationalise that configuration based on the properties of the species from which it may or may not have originated. To understand why Ti^{3+} is $3d^1$ and not $4s^1$, we must compare the stabilities of those two configurations in Ti^{3+}, not the stabilities of orbitals in the neutral atom, which are irrelevant. A critical point that is often missed in the discussion of $4s$ *vs* $3d$ is that the energy of an orbital depends on both n and Z^{eff} ($E = -13.6\left(Z^{eff}/n\right)^2$) and so we need to take both into account (this complication did not arise in the discussion of $2s$ *vs* $2p$ because the two orbitals share the same principal quantum number). The lower value of n for the $3d$ orbital means that if Z^{eff} happens to be equal for a $4s$ and a $3d$ electron (as it is in any hydrogenic atom/ion), it will always be favourable to place the electron in $3d$ rather than $4s$: this is the origin of the 'natural' order that appears in hydrogen. The critical question in poly-electronic atoms is therefore not simply whether Z^{eff} for $4s$ is greater than for $3d$ (because it always is!) but rather whether the difference is great enough to overturn the natural preference for occupation of $3d$ imposed by the difference in n. Let us formulate this problem by asking the following simple question: under what circumstances do the energies of the $3d$ and $4s$ electrons become identical? *i.e.* when does the following equality hold?

$$\frac{Z^{eff}(4s)}{4} = \frac{Z^{eff}(3d)}{3}$$

Clearly it holds when the ratio of effective nuclear charges is

$$\frac{Z^{eff}(4s)}{Z^{eff}(3d)} = \frac{4}{3}$$

In cases where this ratio is less than 4/3 (in hydrogen, for example, where it is rigorously 1), $3d$ will lie below $4s$ while in cases where it exceeds 4/3, the order will be reversed and $4s$ will fall below $3d$. Let us put some flesh on the bones of this idea by using Slater's rules to calculate Z^{eff} (these are, admittedly, crude, but as we will see, they bring out all of the key factors that are relevant to the discussion). Consider atomic K, and ask whether the 19th electron will prefer to occupy $4s$ or $3d$. We know that the correct answer is $4s$, so we should anticipate that the ratio $\frac{Z^{eff}(4s)}{Z^{eff}(3d)}$ will exceed the critical value of 4/3.

For $1s^2 2s^2 2p^6 3s^2 3p^6 4s^1$: $\quad Z^{eff}(4s) = 19 - 10 \times 1 - 8 \times (0.85) = 19 - 16.8 = 2.2$

For $1s^2 2s^2 2p^6 3s^2 3p^6 3d^1$: $\quad Z^{eff}(3d) = 19 - 10 \times 1 - 8 \times (1.0) = 19 - 18 = 1.0$

$$\frac{Z^{eff}(4s)}{Z^{eff}(3d)} = \frac{2.2}{1.0} = 2.2 > \frac{4}{3}$$

and so we confirm that the difference in Z^{eff} for the two electrons is indeed sufficiently large to overturn the preference for occupation of $3d$ imposed by the principal quantum number, n. If we follow this analysis through and ask where, in the periodic table, the crossover point occurs where (core)$4s^1$ becomes favoured over (core)$3d^1$, we find that it is at Si, where Z^{eff} (4s)=1.45 and $Z^{eff}(3d) = 1.0$. This is somewhat later than is observed experimentally - the excited ${}^4P$ ($1s^2 2s^2 2p^4 s^1$) state of nitrogen is already more stable than ${}^2P$ ($1s^2 2s^2 2p^2 3d^1$)

- probably due to the approximations inherent in Slater's rules. Nevertheless, the point is clear that the crossover occurs well before K, which is the focus of attention in this discussion simply because it is lightest element where the $4s/3d$ crossover is reflected in the ground-state configuration rather than an excited state. What happens, then, if we consider a transition metal cation with the same number of electrons as K – Ti^{3+} or V^{4+}, for example. The rule of thumb is that when we ionise a transition metal, we remove the $4s$ electrons 'first' (but note the earlier comments about this phraseology being unhelpful!), so the configuration will be $1s^2 2s^2 2p^6 3s^2 3p^6 3d^1$ and not $1s^2 2s^2 2p^6 3s^2 3p^6 4s^1$, as it was in K. Let us follow the same argument through, except now use Ti^{3+} rather than K. The only difference is that the nuclear charge has increased by 3 units: everything else is identical. The effective nuclear charges for the two configurations are therefore:

$1s^2 2s^2 2p^6 3s^2 3p^6 4s^1$: $\quad Z^{eff}(4s) = 22 - 10 \times 1 - 8 \times (0.85) = 22 - 16.8 = 5.2$

$1s^2 2s^2 2p^6 3s^2 3p^6 3d^1$: $\quad Z^{eff}(3d) = 22 - 10 \times 1 - 8 \times (1.0) = 22 - 18 = 4.0$

$$\frac{Z^{eff}(4s)}{Z^{eff}(3d)} = \frac{5.2}{4.0} = 1.3 < \frac{4}{3}$$

has now fallen below the critical value of 4/3. The difference in Z^{eff} is therefore no longer sufficient to overturn the difference in n and we revert to the normal hydrogenic ordering. Notice that $Z^{eff}(4s)$ is still greater than $Z^{eff}(3d)$, it is just that the ratio of the two is no longer sufficient to overturn the difference in n. The preference for $3d^1$ for transition metal ions simply reflects the fact that these ions are more hydrogenic than atomic K is (by 'more hydrogenic', I mean that the ratio $(Z^{eff}(4s))/(Z^{eff}(3d))$ is closer to one). Given the approximations in Slater's rules, one might argue that the value of 1.3 is only marginally less than 4/3 and therefore that this conclusion is unsound. We can, however, generalise the argument to any atom/ion M^{q+} with 19 electrons (K, Ca^{+}, Sc^{2+}, Ti^{3+}, V^{4+} *etc etc*), where the ratio of effective nuclear charges is given by:

$$\frac{Z^{eff}(4s)}{Z^{eff}(3d)} = \frac{Z-16.8}{Z-18} = \frac{19+q-16.8}{19+q-18} = \frac{q+2.2}{q+1}$$

Clearly this ratio tends towards a limiting value of 1.0 as q becomes very large – in other words, the atom/ion becomes increasingly more hydrogenic as the positive charge on the nucleus increases, and this conclusion is entirely independent of the precise values chosen for the shielding constants. Moreover, if we change the number of electrons (comparing $4s^1 3d^1$ *vs* $3d^2$, for example), the constants (2.2 and 1 in the equation above) will take different values, but this doesn't alter the fact that the ratio $(Z^{eff}(4s))/(Z^{eff}(3d))$ tends to 1 in the limit of large q. Thus, the argument is entirely general: irrespective of the d electron count, the electrons will always end up in the $3d$ orbitals rather than $4s$ in highly charged species. Thus far the key feature of the argument for the preference for the $4s^1$ configuration over $3d^1$ in K is that Z^{eff} is significantly greater in the former, but we have not yet addressed the critical question of why Z^{eff} is greater for $4s$ than for $3d$. The classic textbook argument is that the $4s$ orbital penetrates the core ($1s^2 2s^2 2p^6 3s^2 3p^6$ for K) more effectively than $3d$, and so 'experiences more of the positive charge' (see the radial distribution functions in Figure 3 below). This seems, at first glance, to be borne out by the larger Z^{eff} for $4s$ (2.2 *vs* 1.0 for $3d$). We can, however, raise all of the same objections to this statement as we did to the equivalent one for $2s$ *vs* $2p$ in Li: by absorbing the physically real electron-electron repulsions into an artificially reduced attraction to the nucleus, we place an undue emphasis on attraction rather than the (now deliberately concealed) repulsions. This, in turn, tempts us to imagine that the only way to increase the stability of an electron (in $3d$ or in $4s$) is to increase the attraction to the nucleus, which fixes our attention on proximity of the electron to the nucleus as the key factor. In the absence of other electrons this would, of course, be entirely reasonable, but we know that in the absence of other electrons the $3d$ orbital is more stable than $4s$, not *vice versa.* This is of course perfectly natural: an electron in the $3d$ orbital is, on average, much closer to the nucleus than is $4s$, and so in the absence of competing repulsions, the greater attraction to the nucleus wins out. Note again here, as in the $2s/2p$ case, that the focus is squarely on the large outer radial maximum where the vast majority of the electron density is, not on the much smaller radial maxima! So what really changes when electrons are added to the core? As we progressively fill the inner shells, the radial maximum of $3d$ is still closer to the nucleus than the outer maximum in $4s$, so the attraction to the nucleus is still stronger in $3d$. However, the radial maximum of $3d$ is now also much closer to the ever-expanding core, and $3d$-core repulsions become particularly acute when we reach the 3rd row and start to fill $3s$ and $3p$, which have radial maxima in the same region of space as $3d$. This is why Slater's rules differentiate $4s$ and $3d$ only through their shielding by the $3s$,$3p$ core shell: each of the $3s$,$3p$ core electrons shields a $3d$ electron by 1.0 unit of charge, but a $4s$ electron by only 0.85 - all other shielding constants are identical. The difference between 0.85 and 1.0 reflects the reality that the $4s$ electron is further away from the $3s$,$3p$ core than is $3d$, and therefore repelled less by it. In short, the real reason for the higher Z^{eff} for the $4s$ electron in K is not a greater attraction to the nucleus, but rather the lower repulsions from the expanding core. The expression $E = -13.6(Z^{eff}/n)^2$ actually reflects the tension between attraction to the nucleus (captured in changes in n) and repulsions from other electrons (captured in variations in Z^{eff}). In the example of K, the presence of 18 core electrons means that wherever we place the 19th electron there has to be a compromise. If it enters the $3d$ orbital, the attraction to the nucleus will still be large ($n = 3$) but the repulsions to the core electrons will be large as well (which is why $Z^{eff} = 1$). If, instead, it occupies the $4s$ electron, some of the attraction to the nucleus will be sacrificed because the radial maximum is further out ($n = 4$), but the repulsion from the core electrons will also be reduced for exactly the same reasons ($Z^{eff} = 2.2$). In the case of K, the repulsion dominates and the $4s^1$ configuration is more stable. As we increase the positive charge on the nucleus, both $4s$ and $3d$ electrons are attracted more strongly, but the effect is amplified for $3d$, simply because its radial maximum is closer to the nucleus. The electron-electron repulsions, in contrast, will remain approximately constant, and the result is that the attractive term becomes increasingly dominant over the more subtle differences in electron-electron repulsion that are captured in the shielding constants and the 19th electron drops into the $3d$ orbital. In the limits of extreme high charges (think of U^{73+} as an extreme example of a 19-electron ion!) the differential repulsions experienced by $4s$ and $3d$ become completely insignificant against the background attraction to the nucleus, and the ion becomes almost perfectly hydrogenic. This is the physical basis behind the argument about the limiting behaviour of

$$\frac{Z^{eff}(4s)}{Z^{eff}(3d)} = \frac{q+2.2}{q+1}$$

given earlier: if q is large, the difference in shielding, 2.2 *vs* 1, becomes insignificant.

An analysis of the $4s/3d$ problem using Hartree-Fock theory.

How can I convince a sceptic that that electron-electron repulsion is really the factor that we need to consider rather than the electron-nucleus attraction that is emphasised in most textbooks? Hartree-Fock theory offers a platform to do this, because it simplifies the Hamiltonian in Equation 3 to a mathematically tractable form without completely sacrificing the electron-electron repulsions, as is done in Equation 4. By considering the interaction of each electron with the averaged charge density distribution of all others, the exact Hamiltonian of Equation 3 becomes separable, and can be recast as a series of one-particle equations, the Hartree-Fock equations (Equation 5), which have been the cornerstone of computational quantum chemistry for the past 50 years. The eigenfunctions of these Hartree-Fock equations are the one-electron wavefunctions ("orbitals") that are ubiquitous in chemistry. Note that all three distinct contributions, the kinetic energy, the electron-nucleus attraction and the electron-electron repulsion are explicitly formulated in the Hamiltonian: the repulsive term $\hat{V}_{ee}$ is simplified, but it is not concealed.

$$\hat{H}_i = -\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i} + \sum_{j\neq i}\frac{1}{r_{ij}} = -\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i} + \sum_{j\neq i}(J_{ij} - K_{ij}) \tag{7}$$

The repulsion of electron i by all other electrons, j, is introduced through the Coulomb and exchange integrals, J and K:

$$\hat{J}_{ij}\psi_i(1) = \int \psi_j(2)\psi_j(2)\frac{1}{r_{ij}}\psi_i(1)\,d\tau \tag{8}$$

$$\hat{K}_{ij}\psi_i(1) = \int \psi_j(2)\psi_i(2)\frac{1}{r_{ij}}\psi_j(1)\,d\tau \tag{9}$$

J can be understood in classical terms as the Coulomb repulsion between two charge clouds. The exchange integral K, in contrast, has no classical analogue: it is a purely quantum mechanical phenomenon and is non-zero only when the two electrons have parallel spin. The total repulsion between any pair of electrons is given by $(J_{ij} - K_{ij})$, so K serves to reduce the repulsion between electrons with parallel spins. The fact that each electron experiences repulsion by the average density due to the others means that it effectively loses sight of the explicit position of the other electrons (the electronic motions are not correlated), with the result that repulsions are somewhat overestimated. Correlation is generally a minor effect in absolute terms, but if our goal is to calculate energies to compare to experiment then its omission is a serious failing, and indeed Hartree-Fock theory is rarely used now for such purposes. It has largely been supplanted by two distinct theoretical tools that do incorporate electron correlation, density functional theory (DFT) and configuration interaction (CI). However, neither of these theories is well suited to the analysis I set out here, where the emphasis is on pairwise interactions between electrons in specific orbitals. In DFT, the central variable is the total density rather than the orbitals, and so the orbital-by-orbital repulsions are completely lost (indeed this is the whole point of DFT, because it simplifies the calculations). At the other extreme, the CI wavefunction is built as a linear combination of several different configurations and so, by construction, we have sacrificed the intuitive picture of electrons residing in a specific subset of orbitals. Thus Hartree-Fock theory, although somewhat dated, is uniquely well suited for the analysis set out below. There is, however, a complication in that Hartree-Fock theory treats occupied and unoccupied orbitals differently. Each orbital (whether occupied or empty) is destabilised by repulsions with all other electrons but electrons do not interact with themselves(!). So for an N-electron system, the empty orbitals experience exactly N repulsive interactions while the occupied orbitals experience only N-1 repulsions. The result is that the energies of occupied and vacant orbitals are fundamentally different things which cannot be compared directly. This clearly presents a problem if we seek to analyse the stabilities of $4s^1 3d^0$ or $3d^1 4s^0$ because whichever configuration we choose to study, one of the orbitals is singly occupied while the other is empty. We can, however, work around this problem by selecting systems where all the orbitals that we wish to compare are equally populated. In the following analysis we choose to compare the energies of the $3d$ and $4s$ orbitals in the ^7S states of Cr, Mn^+ and Fe^{2+}, where all orbitals of interest are now singly occupied ($4s^1 3d^5$). If our previous analysis using Z^{eff} works, we should find that the $3d$ orbitals should be progressively stabilised relative to $4s$ as the positive charge increases, and, through this analysis, we hope to discover why this is. Table 1 collects together the total energies (E_{tot}) of the $4s$ and $3d$ orbitals, along with its components, the kinetic (T), electron-nucleus attraction (V_{ne}) and electron-electron repulsion (V_{ee}) terms (in atomic units). The total orbital energies confirm that the Hartree-Fock calculations capture the essence trends: the $3d$ orbital is stabilised relative to $4s$ by 0.3 au for each increase in positive charge. Considering first the data for Cr, it is immediately striking that the near degeneracy of the $4s$ and $3d$ orbitals comes about as a result of the cancellation of very different component parts, T, V_{ne} and V_{ee}. For the $4s$ electron, a rather weak attraction to the nucleus (V_{ne} = –8.29 au) is offset by equally weak electron-electron repulsions (V_{ee} = 7.44 au) and a small kinetic energy term (T = 0.63 au). For a $3d$ electron, in contrast, the attraction to the nucleus is three times larger (V_{ne} = –24.13 au), but so too is the sum of the electron-electron repulsion and kinetic energy (V_{ee} = 18.86 au, T = 4.91 au). After all these terms are added up, the difference in total energies between $4s$ and $3d$ is only 0.14 au. Note that in no sense does the $4s$ electron 'experience a greater attraction to the nucleus' than $3d$ - V_{ne} is three times larger for $3d$, simply because $3d$ is much closer to the nucleus!

Table 1: Components of the self-consistent energies of the $4s$ and $3d$ orbitals of the ^7S states of Cr, Mn^+ and Fe^{2+} (in au) at the HF/VTZ level.

	T	V_{ne}	V_{ee}	E_{tot}
		Cr		
$4s$	0.63	–8.29	7.44	–0.22
$3d$	4.91	–24.13	18.86	–0.36
		Mn^+		
$4s$	1.13	–10.96	9.28	–0.54
$3d$	6.50	–29.26	21.84	–0.93
		Fe^{2+}		
$4s$	1.71	–13.52	10.86	–0.95
$3d$	8.13	–34.25	24.47	–1.65

Hartree-Fock theory allows us to go a step further and decompose the total electron-electron repulsion into separate components for each pair of electrons, which are shown in Figure 4. As noted above, the total V_{ee} is very much greater for a $3d$ electron (+18.86 au) than for a $4s$ electron (+7.44 au), and the major contribution to this difference comes from repulsions to the electrons in the $3s$,$3p$ core, which have radial maxima in the same region as $3d$. So the repulsion between a $3d$ electron and each $3p$ electron ($V_{(3d-3p)}$) costs 0.72 au while the repulsion between the $4s$ electron and each $3p$ electron ($V_{(4s-3p)}$) costs only 0.32 au. This is the physical phenomenon that underpins the difference between Slater's shielding constants (1.0 for $3d$ but only 0.85 for $4s$) and, ultimately therefore, the differences in Z^{eff}. Note once more the subtle change in emphasis: that the $4s$ electron experiences a higher effective nuclear charge than $3d$ is not in doubt, but it is not because the inner radial maxima of $4s$ allows the electron to access the nucleus – rather it is because the very large and diffuse outer radial maximum allows it to avoid all the other electrons. Given the huge variations in the component parts of the total energy, it should probably come as no surprise that it has proven so difficult to explain trends in the relative order of $3d$ and $4s$ in terms of a single 'catch-all' parameter, Z^{eff}.

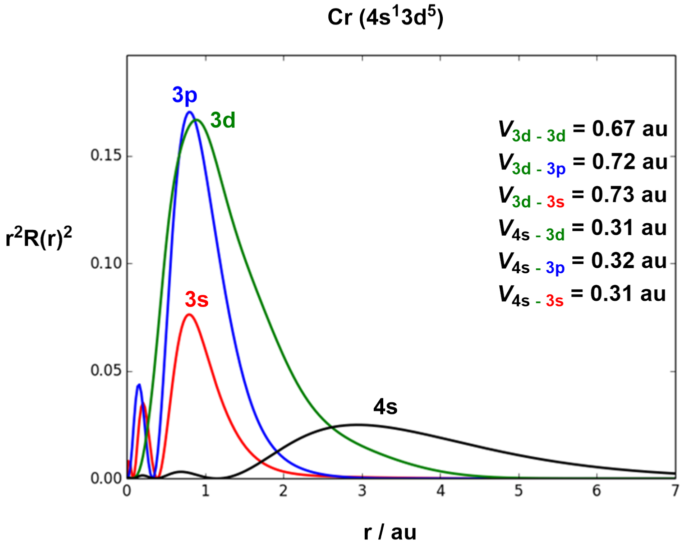

Figure 4: Components of the electron-electron repulsion in the ^7S state of Cr.)

If we now look at how the various components of the orbital energies depend on the charge on the ion, we note that the overall 0.3 au relative stabilisation of the $3d$ orbitals per unit increase in charge is also a result of compensation between much larger fluctuations in the individual components of the energy. The magnitudes of all three components in Table 1, T, V_{ne} and V_{ee}, inevitably increase with charge as the orbitals become more compact. For an electron in the $3d$ orbital, a unit increase in charge causes a 5.1 au increase in $|V_{ne}|$ set against 2.8 au increase in V_{ee} and 1.6 au increase in T. For an electron in the $4s$ orbital, in contrast, all three components are affected to a lesser degree: $|V_{ne}|$ increases by only 2.6 au while V_{ee} and T increase by 1.8 au and 0.5 au, respectively. The dominant factor in stabilising the $3d$ orbital over $4s$ is therefore the very dramatic increase in $V_{ne} = \langle \frac{Z}{r_i} \rangle$ experienced by the $3d$ electron: this is, again, simple electrostatics - the $3d$ electron is closer to the nucleus than $4s$, so is stabilised more by an increase in charge, $V_i = (q_i q_j)/r_{ij}$. The electron-electron repulsion, V_{ee}, changes much less down the series because the number of electrons remains the same – the small increase is a secondary consequence of the contraction of the orbitals.

The result is, therefore, that the delicate balance between attraction and repulsion is tipped progressively in favour of the former as the positive charge increases, leading to occupation of the 3*d* orbital. This is the physical basis for the argument that the hydrogenic order will always be restored in the limit of high charge on the ion. A perspective on the Madelung order and the periodic table The balance between attraction to the nucleus and repulsion with other electrons that lies at the heart of this article can be used to understand the so-called 'Madelung order' that determines the shape of the periodic table. The Madelung rule states that orbitals are filled in order of increasing (n + l) and, in cases where (n + l) is equal, the orbital with higher n is filled first.

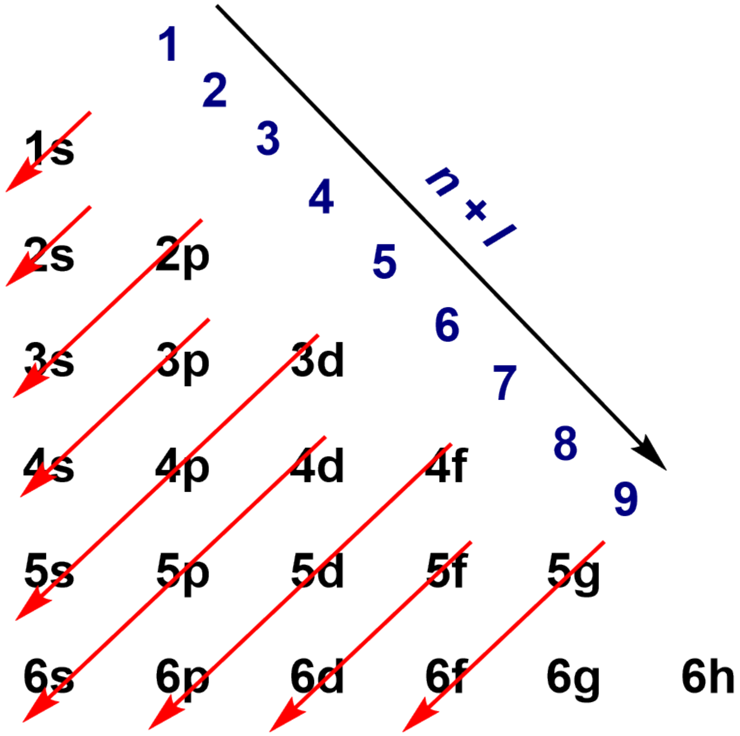

Figure 5: The Madelung order of filling in the periodic table.)

We have already discussed the 2*s*/2*p* balance, and concluded that the greater repulsion with the $1s^2$ core incurred by the 2*p* electrons dictates that it is the 2*s* orbital (n + l = 2) that is occupied by the third electron in Li and not 2*p* (n + l = 1). Likewise we have seen that when we reach the start of the 4th period, the 19th electron enters the 4*s* orbital in preference to 3*d* because the former incurs less repulsion from the core (and in particular from the $3s^2 3p^6$ shell). It is often emphasised that the Madelung order refers only to neutral atoms – as we have seen, in the cations the differential shielding between $4s$ and $3d$ becomes less important compared to the attraction to the nucleus, the ion becomes increasing hydrogenic, and the order of filling is determined first by increasing n, and then by increasing l. Even for the neutral transition metal atoms, however, there are a number of well-documented exceptions to the Madelung rule, the most well-known of which is Cr, where the ground-state configuration is $4s^1 3d^5$ and not $4s^2 3d^4$. This reflects the subtle re-balancing of attraction and repulsion associated with the half-filled $3d$ sub-shell: the exchange interaction between electrons of the same spin means that a 5th electron can enter the $3d$ sub-shell without incurring as much repulsion (primarily to the other $3d$ electrons) as a 6th or 7th electron would, and as a result the increased attraction to the nucleus available in the $3d$ orbital is sufficient to tip the balance in favour of $4s^1$3d^5. The next key 'decision point' in the periodic table comes at the start of the 6th period at element 55 (Cs), at which point the 54-electron core contains $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 5s^2 5p^6$. The 55th electron could then, in principle, occupy any one of the 6s, $4f$ or $5d$ orbitals, and the Madelung order determines that the first of these is preferred and indeed Cs has the $6s^1$ configuration. The energetic balance between $6s^1$, $5d^1$ and $4f^1$ is determined by precisely the same factors as the $4s^1/3d^1$ balance: the $4f^1$ configuration maximises attractions to the nucleus because the radial maximum is close, but also maximises the repulsions with the core electrons. The $6s^1$ configuration, in contrast, minimises repulsions with the core but only at the expense of sacrificing much of the attraction to the nucleus. The $5d^1$ configuration represents a compromise situation, where the attraction to the nucleus is intermediate (less than for $4f$ but greater than for $6s$), as is the repulsion from the core (again, less than for $4f^1$ but more than for $6s^1$). The total energies of the three configurations are then determined by the sum of attractive and repulsive terms that differ widely, and it is no surprise that the identity of the most stable one is hard to predict. It is also no surprisethat there are exceptions to the Madelung order as we cross the lanthanide series - the ground state configuration of

La, for example, is $6s^2 5d^1$ and not $6s^2 4f^1$, as the Madelung order would demand and Gd is $6s^2 5d^{14} f^7$ rather than $6s^2 4f^8$. In the latter case, the very strong repulsions experienced by an 8th electron in the 4f shell tips the balance in favour of occupying the $5d$, despite the necessary sacrifice in attraction to the nucleus. As we increase the positive charge, however, all of the subtle differences in the repulsions experienced by $6s/5d/4f$ electrons become insignificant relative to the attraction to the increasing positive charge at the nucleus, and so the hydrogenic order is restored and all cations have the $4f^n$ configuration.

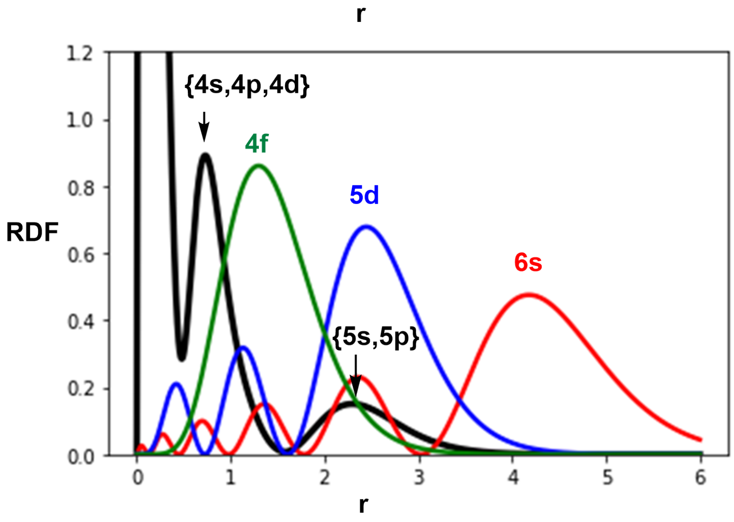

Figure 6: Spatial arrangement of valence and core orbitals at the start of the 6th period.)

Summary

The concept of an effective nuclear charge is without doubt a powerful one, but by concealing the mathematically difficult electron-electron repulsions behind the cloak of a reduced nuclear charge, we promote the unbalanced view that stability is solely a matter of maximising attractive forces. In fact, stability is always a compromise between attraction and repulsion, and can therefore be increased either by increasing the former or by reducing the latter. From this rather different perspective, we see that the stability of the $1s^2 2s^1$ configuration relative to $1s^2 2p^1$ in Li arises not because the inner radial maximum of the $2s$ orbital enhances interactions with the nucleus, but rather because its diffuse outer radial maximum minimises repulsions with the $1s^2$ core. The same arguments provide a coherent picture of the controversial $4s/3d$ crossover at the start of the first transition series. Placing the 19th electron of K or Ti^{3+} in the $3d$ orbital maximises the attraction to the nucleus, but incurs strong repulsions with the $3s$ and $3p$ core electrons. Occupying the $4s$ orbital, in contrast, sacrifices much of the Coulomb attraction to the nucleus, but also attenuates much of the repulsion with the core electrons. Which of these factors wins out depends on the system: the larger nuclear charge in cations amplifies the attractive term, favouring the $3d^1$ configuration, while electron-electron repulsions hold sway in neutral species such as K, driving the adoption of a $4s^1$ configuration. The Madelung order that underpins the structure of the periodic table can also be understood in these terms. The principal quantum number determines, to a very good approximation, the position of the outer (and largest) maximum in the radial distribution function. In hydrogen, where attraction to the nucleus is all that matters, orbitals with lowest principal quantum number are filled first. As we move across the periodic table, however, orbitals with lower principal quantum numbers suffer more repulsions from the ever-expanding core than those with higher principal quantum numbers, simply because they are closer to it. There are, therefore, inevitably crossover points where orbitals or different quantum numbers are near degenerate as a result of the balance between attraction and repulsion. Classic cases occur at the start of the 4th and 5th periods, where nd (strong attraction, strong repulsion) and (n+1)s (weak attraction, weak repulsion) are almost degenerate and also at the start of the 6th and 7th periods, where nf (strong attraction, strong repulsion), $(n+1)d$ (intermediate attraction, intermediate repulsion) and $(n+2)s$ (weak attraction, weak repulsion) are close.

Methods

The Hartree-Fock calculations and the decomposition of the electron-electron repulsion terms were performed using a locally developed python programme, Oxelec. The code is available at http://mcgrady.chem.ox.ac.uk/resources.html. For the calculations on Ne the uncontracted (23s16p) Universal Gaussian Basis Set of Jorge et al. was used as a high degree of flexibility was required to describe both the neutral atom and the highly charged hydrogenic reference states. For the comparison of 4*s* and 3*d* in the transition metals Ahlrich's valence triple zeta (VTZ) basis set was used.

Appendix: analysis of the origins of the $2s/2p$ ordering in the second period using Hartree-Fock theory.

Ideally, we would approach this problem by assessing the relative stabilities of the 2*s* and 2*p* orbitals in Li, where the difference is reflected in the ground-state configuration. However, the Hartree-Fock methodology gives an unbalanced treatment of the energies of occupied and virtual orbitals (as noted in the 4*s*/3*d* discussion) which complicates the comparison when one orbital is filled and the other empty. We therefore recast the problem by focussing on Ne, where all orbitals (2*s* and 2*p*) are equally occupied, and we ask the simple question: why is the 2*s* orbital more stable than 2*p* in this case? The ^2S ionisation in the photoelectron spectrum of Ne lies 29 eV below ^2P, so the difference in orbitals energies is clearly very pronounced. The final results of the Hartree-Fock calculation, labelled "self-consistent orbitals", are summarised at the bottom of Table 2, where the 2*s* orbital is indeed more stable than 2*p*: the energies (E_{tot}) are –1.93 au and –0.85 au respectively, a difference of 29.4 eV. The calculations seem to reproduce the experimentally observations more than adequately, offering some encouragement that the decomposition of the orbital energies into their component parts (kinetic energy, T, attraction to the nucleus, V_{ne} and electron-electron repulsion, V_{ee}) will be meaningful. The data in Table 2 suggest, at first glance, that the greater stability of the 2*s* orbital can be traced primarily to a stronger interaction with the nucleus: $V_{ne} =$ –16.33 au for 2*s* *vs* –14.35 au for 2*p*), apparently consistent with the textbook explanation that the 2*s* electron 'experiences more of the nuclear charge'. There is, however, a twist in the tale. In textbook explanations, the orbitals are generally assumed to adopt their hydrogenic shapes (or at least it is never explicitly stated that they do not!). The electrostatic interaction between a nucleus, A, and a charge distribution represented by $\psi^*\psi$ is given by the expectation value:

$$V_{ne} = \langle\psi|\frac{Z_A}{r_1 2}|\psi\rangle = Z_A\,\langle\frac{1}{r_{12}}\rangle$$

which, in a hydrogenic atom or ion, depends only on the principal quantum number n and not on l (its value is in fact Z_A/n^2). The very fact that V_{ne} (and indeed T) are not identical for 2*s* and 2*p* therefore immediately alerts us to the fact that two orbitals must deviate from their hydrogenic forms. So if we want to argue that "greater attraction to the nucleus" is behind the greater stability of 2*s* *vs* 2*p*, we have to address the question of why and how the orbitals deviate from their hydrogenic forms. A logical starting point is therefore to ask how the total energy and its various components (V_{ne}, V_{ee} and T) would behave if the orbitals were forced to retain their hydrogenic form (which, after all, is the unstated assumption in most textbooks). In Figure A1(a) radial distribution functions for the hydrogenic 1s, 2*s* and 2*p* orbitals are shown, calculated using a positive charge of Z = +7.5 at the nucleus ("hydrogenic" means that there are no electron-electron repulsions). This choice of Z requires further comment: it was selected because it generates 2*s* and 2*p* orbitals with radial maxima at approximately the same positions as those in the self-consistent solutions (vide infra). In contrast if we assume perfect screening (Z = +1) or zero screening (Z = +10), the resultant orbitals are much more diffuse or much more contracted, respectively, than the self-consistent ones. The choice of Z = 7.5 therefore a reasonable compromise, but the key conclusions that drawn below do not depend in any significant way on the precise value.

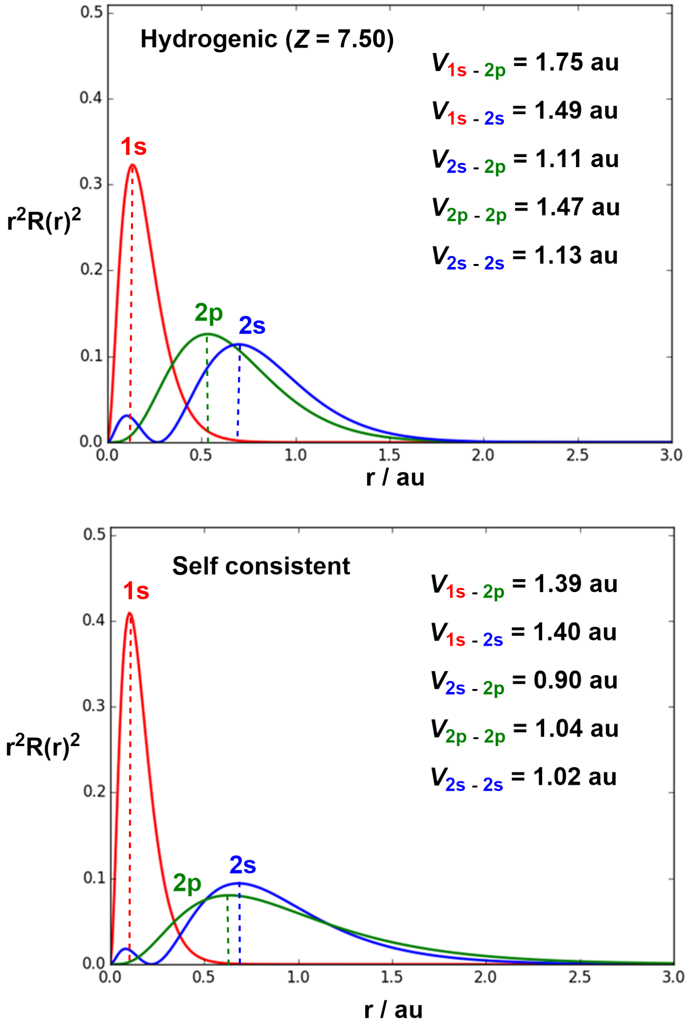

Figure 7: Radial distribution functions and components of the electron-electron repulsion in the Ne atom (a) in the hydrogenic limit (Z = +7.50) and (b) at self-consistency.)

The total energies of the $2s$ and $2p$ electrons in the Z = 7.50 hydrogenic case (Table 2, top) are identical (= –7.03 au), as are their components T (= +7.03 au) and V_{ne} (= –14.06 au), as they must be because neither depends on l. If the 10 electrons are then added without allowing the orbitals to relax (middle entry in Table 2, "frozen orbitals" and also Figure A1(a)) the interaction with the nucleus (V_{ne}) remains absolutely unchanged and identical for $2s$ and $2p$, as it must do according to equation 6. Nevertheless, we immediately recover the correct ordering of the total energies, $2s < 2p$, and in fact the difference of 1.34 au is even larger than in the self-consistent solution. However, the difference is now contained entirely in the electron-electron repulsion term, which is higher for a $2p$ electron than its $2s$ counterpart – this is an inevitable conclusion, because V_{ne} and T cannot change. The more punitive electron-electron repulsion experienced by a $2p$ electron is a cumulative effect of interactions with the two electrons in the 1s core (V_{1s-2p}= 1.75 au), with the two electrons in the $2s$ orbital (V_{2s-2p} = 1.11 au) and, most damagingly, with the other five $2p$ electrons (V_{2p-2p}= 1.47 au). The corresponding numbers for an electron in the $2s$ orbital are all significantly lower: (V_{1s-2s}= 1.49 au, V_{2s-2s} = 1.13 au). These differences can all be understood by reference to the radial distribution functions in Figure 1. The $2p$ orbital has only a single radial maximum, which means that all the $2p$ electrons are confined to a narrow shell. The $2s$ orbital, in contrast, extends across a much wider range of radii on either side of its node, with the result that two $2s$ electrons are, on average, further apart from each other than are two $2p$ electrons. The $2p$-$2p$ repulsions are therefore much worse than $2s$-$2s$. Generalising this argument, electron-electron repulsions will always be strongest in orbitals that lack a radial node, i.e. $2p$, $3d$ and 4f, and many of the unique properties of the first transition series and the lanthanides can be traced to precisely this feature. The presence of a radial node in the $2s$ orbital places the outer (and much larger) radial maximum at a greater distance from the $1s^2$ core than the radial maximum for $2p$, with the critical result that V_{1s-2s} is lower than V_{1s-2p}. The key point here is that it is the position of the outer radial maximum in $2s$ that really matters here, not the inner one.

Textbook treatments do not usually extend to the possibility that orbitals may not be hydrogenic, but let us explore what changes when the orbitals are allowed to relax to their self-consistent form. The transition from the hydrogenic situation to self-consistency represents a compromise between kinetic energy, attraction and repulsion: the algorithm seeks to relax the orbitals in such a way that the attraction to the nucleus is maximised while the kinetic energy and electron-electron repulsions are minimised. This is necessarily a compromise because a build-up of electron density near the nucleus maximises the Coulomb attraction to the nucleus, but also the maximises repulsion between the electrons. Conversely

Table 2: Components of the self-consistent energies of the $2s$ and $2p$ orbitals of Ne from Hartree-Fock theory (all values in atomic units)

	T	V_{ne}	V_{ee}	E_{tot}
Hydrogenic (Z = 7.50)				
$2s$	7.03	–14.06		–7.03
$2p$	7.03	–14.06		–7.03
Orbitals frozen in hydrogenic form				
$2s$	7.03	–14.06	10.75	+3.72
$2p$	7.03	–14.06	12.27	+5.24
Self-consistent orbitals				
$2s$	5.21	–16.33	9.18	–1.93
$2p$	4.26	–14.35	9.24	–0.85

a more diffuse electron density distribution alleviates some of the electron-electron repulsion but only at the expense of weaker attraction to the nucleus. The orbitals which deviate most from their hydrogenic form will be those that experience the most punitive electron-electron repulsions in the first place – i.e. the $2p$ orbitals (V_{2p-2p}= 1.47 au *vs* V_{2s-2s}= 1.13 au). These $2p$-$2p$ repulsions are indeed much reduced in the self-consistent orbitals (V_{2p-2p}= 1.04 au), where the maximum in the $2p$ radial distribution function has become much broader and more diffuse. Inevitably, this also means that V_{ne} is less attractive and the kinetic energy, T, is substantially reduced. The $2s$ radial distribution function, in contrast, rather similar in the hydrogenic and self-consistent cases, simply because the $2s$ electron experienced less punitive electron-electron repulsions in the first place: V_{2s-2s}, for example, remains relatively unchanged from its hydrogenic value (1.02 au *vs* 1.13 au). So finally we have all the necessary information to answer the question of why the $2s$ electron is more stable than $2p$ in Ne. If we start from a hydrogenic reference point, we see that $2p$ electrons are repelled by the other electrons to a far greater extent than are the $2s$. As a result, the self-consistent $2p$ orbitals become more diffuse in order to alleviate the repulsion, which has the side effect of reducing the interaction with the nucleus. Nevertheless, the sum of reduced repulsion and reduced attraction results in a net stabilisation. The $2s$ orbitals, in contrast, deviate less from their hydrogenic starting point simply because they are naturally more diffuse to start with, and are therefore less strongly repelled by other electrons. In the final (self-consistent) solution, the repulsions experienced by an electron in $2s$ or $2p$ are approximately equalised, but the $2p$ electron has had to sacrifice a greater fraction of its attraction to the nucleus in order to achieve this. In short, V_{ne} is indeed more attractive for a $2s$ electron compared to $2p$, but it is not because the $2s$ electron is pulled towards the nucleus (compared to the hydrogenic reference), it is because the $2p$ electron is pushed away from the $1s^2$ core! So with the emphasis now firmly on electron-electron repulsion as the root cause of orbital stability, where does 'penetration' come into the argument? The number or radial nodes an orbital has clearly correlates with stability: $2s$ has one more node than $2p$ and it is indisputably more stable, but it is the outer maximum in the $2s$ radial distribution function that matters: it lies further from the nucleus, allowing it to better avoid the other electrons. Although we cannot address directly the question of the preference of the lithium atom for the $1s^2 2s^1$ configuration over $1s^2 2p^1$ using Hartree-Fock theory, we can extrapolate the arguments made above for Ne . The V_{2s-2s}, V_{2s-2p} and V_{2p-2p} components of the repulsion are clearly irrelevant in this context because the $2s$,2p manifold contains only one electron. The greater stability of the $2s$ orbital must, therefore, have its origins in different repulsions with the electrons in the $1s^2$ core. The smaller value of V_{1s-2s} compared to V_{1s-2p} is therefore the key factor, forcing the $2p$ electron to sacrifice more of its Coulomb attraction to the nucleus to alleviate the punitive electron-electron repulsions at self-consistency. Again note the change in emphasis: the $2s$ orbital is certainly stabilised relative to $2p$, but not because the small inner radial maximum of $2s$ affords better attraction to the nucleus - rather it is because the diffuse outer radial maximum reduces the repulsions to the $1s^2$ core electrons.