



Universiteit
Leiden
The Netherlands

Holography, fermi surfaces and criticality

Čubrović, M.

Citation

Čubrović, M. (2013, February 27). *Holography, fermi surfaces and criticality. Casimir PhD Series*. Retrieved from <https://hdl.handle.net/1887/20607>

Version: Not Applicable (or Unknown)
License: [Leiden University Non-exclusive license](#)
Downloaded from: <https://hdl.handle.net/1887/20607>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/20607> holds various files of this Leiden University dissertation.

Author: Čubrović, Mihailo

Title: Holography, Fermi surfaces and criticality

Issue Date: 2013-02-27

Chapter 1

Introduction

1.1 Holographic principle: the idea

Reductionism lies at the heart of physics. Much of the history of physics can be understood as striving for reduction in the number of basic principles and thus explaining seemingly disparate phenomena starting from the same core idea. Indeed some of the key scientific revolutions can be formulated in terms of unifying previously distinct areas of study: Newtonian mechanics bridges the gap between statics and dynamics, Maxwell electrodynamics connects electricity and magnetism, Boltzmann's kinetics unites mechanics and statistical physics. General relativity has unified gravity with mechanics while quantum field theory brought a unified look at quantum mechanics, electrodynamics and statistical physics. Finally, in the last decades we are witnessing the attempts at unifying all of physics within string theory. Looking for analogies between different systems has certainly proven to be one of the deepest principles in the search for fundamental laws of nature.

The presumed approach of the Theory of Everything through the advent of string theory (if it indeed turns out to lead to the Theory of Everything) in parallel with the standing fundamental problems of many-body and collective physics – such as unconventional superconductivity and quark confinement – has actualized the problem of emergence versus reductionism. We are facing the question of how the reduction to the few fundamental principles might help us with resolving the problems which obviously come from a complicated interplay of an enormous number of degrees of freedom. One could even wonder if extremely complex systems

are within the reach of microscopic models at all – after all, we know that hydrodynamics is not within the reach of the single-molecule description. Such a question, in its full generality, is hard to address, and the answer almost certainly varies – systems which do not at all have a single dominant energy scale might well be out of reach. On the other hand, successful explanations of collective phenomena such as Mott insulators, or the energy cascade in turbulence do give a hint that reduction to the basic principles can be fruitful even if these principles live on the scales which are many orders of magnitude smaller.

All of the above prompts us to rethink the quest for reduction and analogies as formulated in the first paragraph. We might look for direct analogies between fundamental and emergent phenomena. If Maxwell’s equations connect the two elementary constituents of electromagnetic interaction, are we able to find a theory which connects a fundamental interaction to an emergent phenomenon? Putting it bluntly, is there an analogy between the simple and the complicated? This thesis is an attempt to contribute to the answer in a specific setting – strongly correlated fermions – where the ”complex” side of the duality is likely unreachable by ”ordinary means”¹ and the fundamental side is a string theory through a mapping known as holography.

Holography is an idea aimed at providing a unified description of quantum mechanics and gravity. It was coined from a disparity between the thermodynamically calculated black hole entropy and the naive guess from dimensional analysis. Understood as information content of a physical system, entropy is expected to be an extensive quantity, proportional to the volume (measure) of the system. Nevertheless, the famous semiclassical Hawking-Bekenstein result for the entropy of a neutral non-rotating (Schwarzschild) black hole [9] predicts it as proportional to the surface

¹It is known [109] that the problem of interacting fermions is NP complete. At this place we briefly remind what this means. A problem is said to belong to the NP class if an algorithm exists which *checks* a proposed solution in polynomial time, but no algorithm is known which *finds* a solution in polynomial time. An example could be an equation such that plugging in a given candidate solution and checking if it satisfies the equation can be done in polynomial time, but no polynomial algorithm is known to compute the solution starting from the equation only. Notice that we do not know if such an algorithm really does not exist, or we are simply unable to find it yet (this question is the famous unsolved $P = NP$ problem). NP complete problems are a subclass of NP problems, such that an algorithm that solves an NP problem polynomially could be modified in a certain way to solve *all* NP problems in polynomial time.

area:

$$S = \frac{Ac^3}{4G\hbar} \quad (1.1)$$

Informally, all information about the black hole is stored on a lower-dimensional object, suggesting that a complete description of the black hole in D dimensions can be obtained by looking at the correctly chosen degrees of freedom on a $D - 1$ -dimensional manifold. This is the logic behind the arguments by 't Hooft [122] and Susskind [107]. The foundation of this principle is that it connects the concept of gravity to the quantum-mechanical concept of entropy as counting the states of the system.

The second, more technical key concept in holography is the idea of dualities, mathematically equivalent but different descriptions of the same phenomena – thus providing a bridge between different formalisms or even altogether different physical systems. The idea of duality can be given a very precise and familiar meaning. Formally, it is just a canonical transformation of the action. Well-known examples are the vortex duality for charged scalar fields and electric-magnetic duality in $U(1)$ gauge theory [71]. In the vortex case, the physical picture is that of changing the viewpoint of what is an elementary excitation. If it is the linearly dispersing plane wave, then the vortices then appear as defects where the phase of the charged field winds for a full circle. But if we dualize, then vortices are the elementary excitations and plane waves are complex vortex combinations. The duality can be captured by a Legendre transformation of the action:

$$S = \int d^3x \partial_\mu \Phi \partial^\mu \Phi \mapsto S_{dual} = \int d^3x (a_\mu a^\mu - \partial_\mu \Phi a^\mu). \quad (1.2)$$

Here, Φ is the charged scalar field which lives in two space dimensions, while its gradient $\partial^\mu \Phi$ maps to the vortex field a^μ . We can thus reexpress the action in terms of a^μ : the physics must remain the same but that does not change the fact that some phenomena are much easier to see in one or in the other language. Similar is the wisdom behind the electric-magnetic duality, where the physical observables, i.e. elements of the field strength tensor, transform into each other, again by adding a bilinear term (linear in both old and new components) to the action.

As an idea which connects quantum theories with gravity, holography finds its natural language in the formalism of string theory, where it arises as a duality transformation of the strings themselves. It is within string theory and M theory that a precise realization of the abstract holographic

principle was found. One reason is simply that it offers a coherent framework in which we can study the gravity at various energy scales— from the low-energy description of general relativity to the nonperturbative regime where the string effects dominate.

In string theory language, the duality becomes the equivalence of the open and closed string descriptions. The higher-dimensional, gravitational system is given by the excitations of the closed string. Its lower-dimensional dual gauge field description is given by the excitations on the open strings. In the next section we will present a more detailed explanation of this construction, known as the *AdS/CFT correspondence* [81, 38, 114]. However, in this introductory chapter we will not assume any prior knowledge of string theory on the side of the reader. We will stay away from extensive use of string-theoretical language and results and formulate AdS/CFT in terms of gauge theory and general relativity, with only qualitative discussion of the underlying specifically stringy constructions (branes, open strings between branes, string dualities, etc).

1.2 Realization: AdS/CFT correspondence

We do not intend to give anything like a comprehensive tutorial on AdS/CFT in this (or any subsequent) chapter, we will merely wet the reader's appetite to look for the original references if interested; most of the thesis can be followed without a detailed understanding of the foundations of AdS/CFT. The first explicit realization is due to Maldacena [81]. Here we have a Type IIB superstring theory in a configuration describing a stack of parallel D3 branes (planar objects extending along three spacetime dimensions) at some distance r from each other. The interbrane distance r also determines the "elastic energy" of the open strings which stretch between the branes and carry the gauge fields from a $U(N)$ multiplet: the energy is proportional to r/α' , where α' is the string tension. Consider now the limit of coincident branes, when $r \rightarrow 0$ but with $r/\alpha' = \text{const.}$. In the closed string description, the metric of a stack of coincident D3 branes factors out into the product of AdS space and a sphere: $\text{AdS}_5 \otimes S_5$. The open string description is a very special, highly symmetric QFT – a conformal field theory (CFT). The idea is that the more restricted and special the field theory, the easier it is to relate it to gravity. This certainly does hold for a conformal field theory (CFT), where the very high symmetry severely constraints behavior of correlation functions. CFT has a central

place in modern high and low energy physics – allowing exact calculation of correlation functions in two dimensions and strong results on RG flow (c -theorem [120]) and scaling [24]. In low energy physics they describe the quantum critical systems [15] lying at the heart of the description of phase transitions and strongly competing interactions. The $\mathcal{N} = 4$ supersymmetric Yang-Mills in four spacetime dimensions is such a CFT – despite the many fields involved, its behavior is simple due to conformality, and it has given us the first example of a holographic duality.

This explicit example allows for a quantitative connection between the gauge theory and the supersymmetric theory in AdS geometry. The connection is provided by the fact that the radius of AdS space is proportional to $(gN)^{1/4}$. The supergravity solution can be trusted if $gN \gg 1$ and $N \gg 1$. Remember that this means that the field theory is strongly coupled *and* can be expanded in the inverse number of colors. The lower-dimensional, field theory side in this and similar (early) setups of AdS/CFT are generically non-Abelian gauge theories, either Yang-Mills or its supersymmetric version, motivating another frequently used name for AdS/CFT: *gauge/gravity duality*.

To turn the above discussion into a precise duality, one needs a relation between the partition functions (on-shell actions) of the gauge theory and supergravity. To this end it is critical to determine the boundary conditions for the supergravity fields living in AdS – when they reach the branes, they are coupled to the fields living on them. This was done in the follow-up work by Gubser, Klebanov and Polyakov [38] as well as Witten [114].

1.2.1 Warmup: symmetries

Let us study the closed string (gravity) side first. The formulation of a field theory on AdS spaces is not quite trivial: AdS geometry possesses some troublesome properties such as closed timelike curves and the existence of a boundary at infinity. Informally, anti de Sitter (AdS) space is an open (hyperbolic) equivalent of the perhaps more familiar de Sitter (dS) space. The latter is the solution to the Einstein equations in the vacuum with a positive cosmological constant [2]. The latter can be thought of as a mysterious form of matter with equation of state $p = -\rho$. It has negative pressure: it expands as it cools down, just like our universe. It is thus a cosmological model in the approximation of “empty Universe” where the presence of matter is negligible and the geometry is dictated by the

cosmological constant. In AdS space, on the other hand, the cosmological constant is negative, i.e. it behaves as (positive) pressure of regular matter. Because the matter is cosmological, it cannot clump and one finds a static (time-independent) solution. The Einstein-Hilbert action that describes the anti de Sitter space in $D + 1$ spacetime dimensions is:²

$$S = \int d^{D+1}x (R - \Lambda) \quad (1.3)$$

where R is the scalar curvature while $\Lambda < 0$. As the only dimensionful factor, Λ can be rescaled at will depending on the choice of the unit of length. By convention, we write $\Lambda = -D(D - 1)/L^2$ where L has the meaning of AdS radius. This means that the solution can be embedded into a $D + 2$ -dimensional flat space as a sphere:

$$t^2 - z^2 - y_i y^i = L^2. \quad (1.4)$$

A natural coordinate patch covers half of the space:

$$ds^2 = \frac{r^2}{L^2} (-dt^2 + dx_i dx^i) + \frac{L^2 dr^2}{r^2}. \quad (1.5)$$

The radial coordinate r stretches from 0, called the interior, to infinity, called the AdS boundary. AdS is the maximally symmetric solution to Einstein equations. An extremely useful way to think about AdS_{D+1} is as a hyperboloid embedded in a $D + 2$ -dimensional flat spacetime with signature $(+, \dots, +, -, -)$. The embedding in a spacetime with $D + 2$ dimensions helps to see that the total geometric symmetry group of AdS space is $SO(D, 2)$. The global and local geometry of AdS space are sketched in Fig. 1.1: what globally looks as the usual double hyperboloid (but in Minkowskian as opposed to Euclidean spacetime) locally becomes a patch of isotropic space of "decreasing size" as we move further and further, until at infinity all lengths scale to zero.

One can now motivate the correspondence starting from the symmetry arguments. It is well known that a CFT in D dimensions (one timelike and $D - 1$ spacelike dimension) also obeys the $SO(D, 2)$ symmetry [24]. Informally, the conformal symmetry is just the symmetry associated to

²In this thesis, unless specified otherwise the dimensionality of spacetime is always $D + 1$, the flat space coordinates in D dimensions are denoted by (t, y_i) ($i = 1 \dots D$) and the metric signature follows the convention $ds^2 = -dt^2 + dy_i dy^i$.

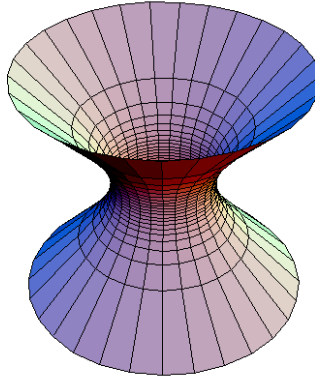


Figure 1.1: Sketch of AdS geometry. Globally it looks like a double hyperboloid but if we take a small patch it becomes very much like Minkowski space in which distances decrease as we move toward infinity. Counterintuitively, local AdS is completely isotropic and has spherical symmetry.

length rescaling, i.e. changing the scale combined with rotations. Conformal field theories are thus closely related to the concepts of self-similarity, fractality and scale-free objects but more general: the scale invariance is continuous, not discrete as in fractals, and it can be broken due to quantum effects – anomalies, like any other physical symmetry. A closer inspection reveals that the exact conformal representation of $SO(D, 2)$ is already geometrically encoded in AdS in a special limit – its boundary transforms in the same way. If one would “glue” some fields on its D -dimensional boundary, these fields ought to be redefinable as representations of the conformal group.

As a result, the CFT can be understood as the boundary degrees of freedom of a field theory in AdS. Emphatically, however, this is *not* enough for a duality, and does not yet encapsulate the idea of AdS/CFT. We need more – not just that AdS space in the near-boundary limit becomes conformal invariant but that *the fields* in AdS in the near-boundary limit also encapsulate the behavior of a conformal field.

1.2.2 Enlightenment: the duality relation

This idea finds its precise formulation in the concept of duality introduced earlier. The quantum theory is dual to gravity, thus *the operators in field*

theory are sourced by the fields on the gravity side. More precisely, the generating functional for the correlation functions in field theory is identified with the minimum of the supergravity action, satisfying specific boundary conditions at the AdS boundary. The precise boundary conditions and the crucial point of AdS/CFT, known as the GKPW (Gubser-Klebanov-Polyakov-Witten) prescription. The prescription addresses the boundary conditions mentioned at the beginning of this section and was proposed in [38, 114]. The conformal and the gravity side are connected through their partition functions as

$$Z_{bnd}(J) = Z_{bulk}(\Phi|_{\partial\text{AdS}} = J) \quad (1.6)$$

where Z_{bnd} and Z_{bulk} are the partition functions on each side, and we have employed Φ as a generic notation for all fields living in the bulk and J are their boundary values, acting as sources. In the classical gravity limit, i.e. for a large N strongly coupled field theory, Z_{bulk} is evaluated simply by plugging in the classical solutions to the equations of motion into the gravity-matter action (in other words, it is the on-shell action). Schematically, this looks like

$$Z_{bulk} = e^{-S(\Phi)}|_{\Phi(r \rightarrow \infty) = J} = \langle e^{\phi J} \rangle_{\text{CFT}} \quad (1.7)$$

where S is the classical gravity action, and in the second equality we have expressed the partition function at the boundary as the generating function of the field theory correlators. The boundary operator ϕ sees the boundary values $\Phi(r \rightarrow \infty) = J$ precisely as sources: treating Z_{bulk} in (1.7) as an effective action for ϕ , we can apply the textbook rule to calculate their correlation functions:

$$\langle \phi(y_1)\phi(y_2)\dots\phi(y_n) \rangle = \lim_{r \rightarrow \infty} \frac{\partial^n e^{-S}}{\partial\Phi(r, y_1)\partial\Phi(r, y_2)\dots\partial\Phi(r, y_n)}$$

This is the essence of applying holography in practice: we do not know how to write Z_{bnd} in terms of boundary fields explicitly, but we can use it as the generating functional of the correlation functions, and thus gain qualitative insight into the system.

The precise translation of the bulk physics into the boundary is thus achieved by analyzing the $r \rightarrow \infty$ limit of various bulk quantities. This is the quantitative basis to constructing the *holographic dictionary* which makes possible numerous practical applications of AdS/CFT. In the next chapter we will introduce dictionary entries such as temperature, chemical

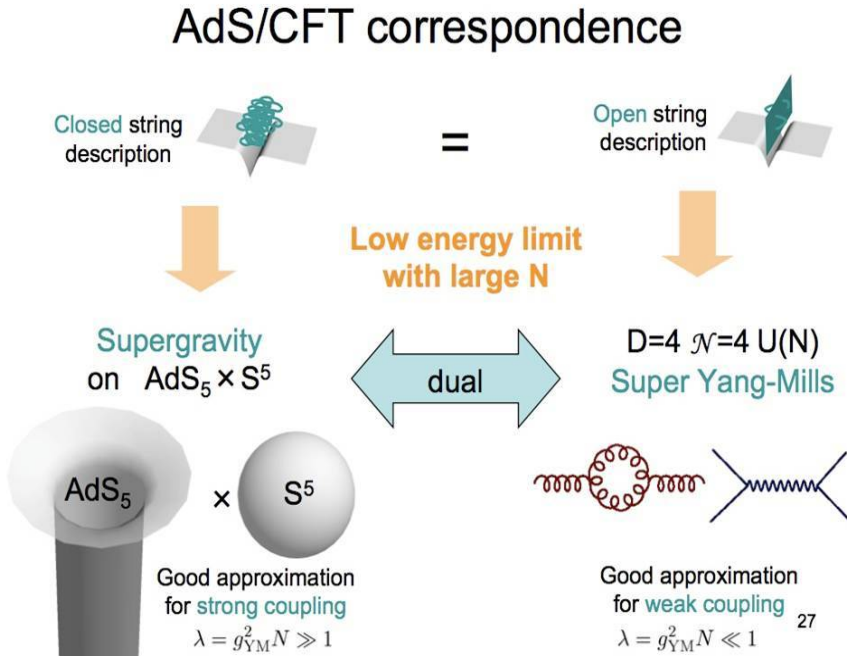


Figure 1.2: Pictorial resume of AdS/CFT: the duality is rooted in the notion of open-closed string duality. On the level of coupling constants it is also a weak-strong duality. Closed string coupling g is related to the coupling g_{YM} of the Yang-Mills theory on open strings as $g_{YM}^2 = g$. The small parameter in the perturbative expansion for the closed string interactions is the combination $\lambda \equiv gN$. Sending the closed string coupling to zero ($g \rightarrow 0$) at constant λ we get classical strings in $AdS_5 \otimes S^5$ while the Yang-Mills theory on open strings reaches the large N limit ($N \rightarrow \infty$). Taking also the limit $\lambda \rightarrow 0$, the classical type IIB string theory becomes type IIB SUGRA.

potential, electromagnetic field, conformal dimension... It is the main link between the formalism of holography and more familiar low-energy QFT physics.

From a more general viewpoint, AdS/CFT was historically important as a facet of the second superstring revolution, which found numerous dualities between string theories with different coupling constants or geometric properties. Here, the control parameter is the combination $1/gN$ of

the string coupling $g = g_{YM}^2$ and the number of colors N . In order to trust the supergravity limit we need $gN \gg 1$, but this is precisely the strongly coupled regime of the gauge theory. Therefore, AdS/CFT is an example of a weak-strong duality. Such dualities are known as S -dualities. Formally, these relate a theory with coupling constant g to a theory with coupling $1/g$. While AdS/CFT does not quite follow this pattern, as the control parameter is not g but gN , it remains a relation between strongly and weakly coupled systems. Needless to say, this gives it a great deal of practical utility: when one side becomes intractable due to string interactions, the other one becomes better and better controlled.

1.2.3 Some general remarks

We will conclude this section with some speculations on broader implications of holography on string theory and other areas. Even though the general holographic principle is essentially a gravity/quantum field theory duality, its full realization in the form of AdS/CFT is a decidedly string-theoretical result, which follows from the near-brane geometry and the action of that solution in a specific brane configuration. In other words, the 't Hooft-Susskind principle states more than AdS/CFT – it states that any physical system with gravity is equivalent to a lower-dimensional system without gravity. One might now wonder if this is indeed so, if holography is in fact a fundamental principle itself, independent of string theory, and a property of gravity and field theory as we know them. There is no answer yet on this central question. At the very least, what one can try is to apply the precise results of AdS/CFT (dictionary entries) to geometries which do not follow from string theory. As long as the geometry looks like AdS at long distances, numerous attempts so far give encouraging results.³ The non-string AdS spaces give us more freedom: we can work in any number of dimensions, with any field content. The price to

³It is much less clear and much more complicated to generalize it to non-AdS spaces, including flat space. This is another important problem to work on. The natural guess is that the correspondence can be generalized to arbitrary geometries and arbitrary field theories. Reasons that require an asymptotically AdS geometry and the difficulties involved in constructing a flat space holography are beyond the scope of this Introduction and indeed this thesis. Roughly speaking, in flat spacetime there seem to be too many degrees of freedom on the gravity side to match to a lower-dimensional QFT; AdS asymptotics puts some rather stringent constraints on the dynamics of gravitational field. The extension beyond AdS is certainly a central fundamental question for the future of holography.

pay is, of course, that we cannot at the present be sure about the consistency of such attempts. This approach is known as *bottom-up* as opposed to the *top-down* string approaches. In this thesis we will mostly use the bottom-up logic, for both practical and conceptual reasons.

In this place it is appropriate to discuss the status of AdS/CFT as a confirmed result versus a conjecture. Though it is widely accepted (e.g. [25]), a rigorous proof is lacking. Nevertheless, the evidence in favor of AdS/CFT is very solid: it has passed numerous non-trivial tests where observables whose forms do not depend on the coupling constant were computed on both sides and compared [2].

1.2.4 Holography outside high-energy theory

The manifestation of holography as a duality has given rise to a completely different research pursuit from the understanding of black holes. Holography can also be used as a tool to understand systems at strong coupling, where the conventional perturbative methods of field theory fail. So far AdS/CFT has established itself as an approach to quantum chromodynamics (QCD) and to condensed matter theory (CMT), the corresponding fields being known as AdS/QCD and AdS/CMT. The power of holography is that it allows us to study previously inaccessible strongly coupled systems. In AdS/QCD, the focus of most work done so far was on describing the confinement transition and studying the quark-gluon plasma at intermediate energies, when neither perturbative QCD nor effective low-energy theories work well (this regime is primarily tested in heavy ion collisions). The latter line of research has produced perhaps the most important result of applied holography so far, the universal viscosity bound, stating that any isotropic equilibrium fluid has an inherent shear viscosity to entropy ratio

$$\frac{\eta}{s} \geq \frac{1}{4\pi} \quad (1.8)$$

The quark-gluon plasma studied in the RHIC accelerator exhibit a viscosity remarkably close to the bound (1.8).

The main approaches exist in AdS/CFT. The first is a top-down approach which constructs a Yang-Mills theory akin to QCD from brane intersections, following closely the early ideas of Witten [114, 115] where the whole endeavor of AdS/CFT is put in the context of specifically gauge/gravity duality, i.e. understanding the dynamics of Yang-Mills fields. The second is a bottom-up scheme where the four-dimensional

QCD is dual an asymptotically AdS₅ space where confinement is modeled by "thinning out" (suppressing exponentially) the amplitudes of fields in the IR. This is done using the insight that the extra dimension in AdS corresponds to the scaling flow in field theory with the near boundary behavior encoding the UV asymptotics. In this case, the RG flow interpretation is that the confinement of low-energy excitations corresponds to suppressing the dynamics in deep interior.

The second claim to fame for AdS/CFT is its application to condensed matter theory. Here, the problems of strong correlations and competing orders show their best (or rather, worst) side. It is thus extremely exciting to see how they dualize in gravity. However, since the phenomenology of condensed matter systems is much richer, and removed even further from the microscopic Hamiltonian, it becomes important to build the model in an appropriate way: to start from the solid and robust features (symmetries, degrees of freedom, extreme limits when some fields decouple or become exactly soluble) rather than engineer the gravity dual in order to get this or another specific phenomenon. The field started with a holographic calculation of transport properties of certain strongly coupled systems [59] and took off with the crucial work of Hartnoll, Horowitz and Herzog on holographic superconductors [47]. Despite the by now universally accepted name, the model in question is not actually a superconductor at all but a boson at finite density which breaks the global phase symmetry by condensing, akin to a superfluid. Nevertheless, precisely as it stands it is a very important proof of concept: this is the simplest possible case of the Landau-Ginzburg picture of order, and thus the obligatory starting point of any candidate theory for description of many-body systems. Holographic superconductors have taken the *bosonic* AdS/CMT to perfection and have been the arena in which many of the universal results and dictionary entries have been obtained.

1.3 The arena: fermions in organized matter

This thesis will focus on AdS/CFT applied to strongly coupled fermion matter. Experimental condensed matter physics has discovered numerous materials which cannot be understood from the weakly coupled perspective. Strongly coupled fermions are thus an experimental reality, and developing general methods to study them is of central importance for understanding the observed phenomena in condensed matter. In AdS/CFT,

precisely the strong coupling regime in field theory is easy to understand on the gravity side, as it corresponds to classical (super)gravity. We will now argue that such holographic description of the strongly coupled physics is especially valuable precisely for fermion systems, as conventional field-theoretical methods are far less helpful for fermions than for bosons.

We have a number of nearly equivalent ways to describe the simple observation that fermions and bosons differ in their behavior. Antisymmetry of fermionic wave functions, the Pauli principle, fermion sign problem and kinematic correlations (i.e., Slater determinants) are all about the fact that the antisymmetry of fermionic states reduces the number of available configurations, acting as a constraint on dynamics and introducing an effective interaction (or correlation) even in absence of any explicit interacting potential. While a non-interacting Fermi gas can still be solved by explicitly taking into account the antisymmetry of states when constructing thermodynamical potentials, presence of interactions spoils the picture: antisymmetry acts as a constraint, and solving an interacting system in the presence of such a constraint becomes a hopeless task. A common way to phrase the problem is the "fermion sign" viewpoint, reviewed e.g. in [119]: it refers to the negative contributions to the fermionic partition function, meaning that it cannot be regarded as a sum of probability amplitudes as for bosons and classical particles.⁴

The fermion signs are simply the minus signs in the density matrix of a fermion system. This is a direct consequence of antisymmetry of the fermionic wave function. For a system of free fermions we can write the wave function exactly; the outcome is the Slater determinant where the odd permutations contribute with a minus sign. Antisymmetry, however, does not depend on interactions in the system and the sign picture will be exactly the same. A technical way to see the trouble with fermion signs is analysis of the fermionic path integral. It is enough to remember the basic rule of constructing the partition function for a system of fermions in compact Euclidean time with period β , thus accounting for finite tem-

⁴Besides condensed matter, another area where the sign problem is well-known is Quantum Chromodynamics (QCD). There, the sign problem arises in a seemingly different but in essence equivalent form: the presence of finite density (and thus chemical potential) makes the Euclidean Hamiltonian non-Hermitian, and thus the partition function complex. The negative vs. complex dichotomy is that of real vs. imaginary time, but in both cases it is the fermionicity of the Hamiltonian which gives rise to problems at finite density, and both negative and complex partition function give us the same pain: absence of probabilistic interpretation.

perature $T = 1/\beta$. Remember that partition function equals the integral of the trace of the density operator:

$$Z = \text{Tr} e^{-\beta H} = \int d^{ND} \mathbf{x} \rho(\mathbf{x}, \mathbf{x}; \beta) \quad (1.9)$$

where the density operator is $\rho(\mathbf{x}_1, \mathbf{x}_2)$ and \mathbf{x} denotes the set of coordinates of all particles in a D -dimensional system with N fermions. Now for a system of indistinguishable particles ρ is a sum over of all permutations Π of the particles, as any two particles can be exchanged without changing the system physically. This gives

$$\rho(\mathbf{x}, \mathbf{x}; \beta) = \frac{1}{N!} \sum_{\Pi} (\pm 1)^{|\Pi|} \rho(\mathbf{x}, \Pi \mathbf{x}; \beta) \quad (1.10)$$

Here, the sum is over all permutations Π of the particles, and $|\Pi|$ is the parity (symmetry/antisymmetry) of the permutation. For bosons all terms are positive and one can define a measure based on the density matrix ρ_+ . For fermions, however, odd permutations carry a negative contribution. The partition function is, of course, always positive, but we see that individual contributions to the density matrix are not. This in turn means that fermions are never classical: unlike for bosons, quantum statistics brings a discontinuity from classical Euclidean field theory and its path integral formulation. The effective action for bosonic expectation values is just the celebrated Ginzburg-Landau theory or one of its many derivations. Nothing like it exists for fermionic operators.⁵ Consequently, despite decades of research of strongly correlated fermions, the actual methodologically sound knowledge we have on this topic is very limited. A measure of the difficulty of the sign problem is the realization of Troyer and Wiese [109] that it is NP complete.

What, then, are the things we do know?

1. *Free Fermi gas*. One example is obvious: the free Fermi gas is exactly soluble. It is not really free, as kinematic correlations are introduced by the statistics, however we know that the Slater determinant accounts for them exactly.

⁵While the expectation value of a fermionic operator is trivially zero, we typically want to compute operator products. Density, correlation functions, transport coefficients etc. are all of this form. However, working with fermion operator products is no easier than working with single fermions.

2. *Fermi liquid.* The second example is the breakthrough of Landau in understanding normal metals in terms of Fermi gases [73]: the Fermi liquid paradigm. The logic is well-known: a gas of particles with infinite lifetimes turns into a gas of quasiparticles with finite but long lifetimes. Everything remains the same as for a free gas, except that all parameters undergo renormalization. The crucial requirement is that the ground state of the interacting system have a nonzero overlap with the ground state in the non-interacting limit. In other words, Fermi liquid is so much akin to a Fermi gas simply because it is adiabatically connected to it. Subsequent, more rigorous studies of the Fermi liquid have confirmed this basic picture (see [5] and references therein). The mathematical foundation of Landau's Fermi liquid insight is provided by the RG formalism for fermions given in [88, 100] and has the form of a functional RG which starts from a weakly interacting theory at intermediate scales and introduces interactions perturbatively in the effective action. Being a weak coupling expansion, it does not have sign problems. However the perturbative treatment does make it hard to treat non-perturbative phenomena e.g. a superconducting instability within this approach.
3. *Fermions in $(1 + 1)d$.* A special case which is in principle completely known is that of fermions in one spatial dimensions. While fashionable these days, and certainly capable of displaying very intricate behavior of correlation functions and transport properties (see e.g. [110]), one-dimensional fermions are completely demystified by bosonization: in one space dimension, any fermion system can be bosonized in infinitely many ways (the most typical situation is the spin-charge separation) and then solved through usual field theory methods. The reason is that statistics cannot really be defined in $1 + 1$ dimension: the manifold of possible Slater determinant states coincides with the manifold of nodeless wave functions.
4. *Miscellanea.* Finally, there is a small number of exactly soluble interacting fermion models in higher dimension, such as exact wave functions for Fractional Quantum Hall states [74]. These are however of very little significance for the broader sign problem, being rather special non-generic.

The inescapable conclusion is that, if we want to avoid the strange ad hoc models of the point (4), everything we know is either to bosonize

or to hope that the system studied is adiabatically connected to a non-interacting Fermi gas, at least in the IR. The vast field of strongly correlated electrons armed with various field-theoretical techniques [34, 110] is as it stands incapable of constructing (through controlled, justifiable approximations) *novel ground states* of fermion matter. The list of celebrated experimental puzzles, from unconventional superconductors [118, 42] to heavy fermions [80], all likely novel ground states qualitatively different from normal Fermi liquids, is therefore in desperate need of a theoretical paradigm that will not depend on non-interacting or bosonic physics.

We are now able to formulate a sharp question underlying all of strongly correlated electron systems: *Is there a stable state of electrons at finite density which cannot be adiabatically continued to Fermi gas?* This is perhaps the closest it comes to formulating the motivation for this thesis in one sentence. A solution we propose here is to use the power of holography.⁶

This is not just an academic question. The importance of strongly correlated electron physics is its manifest necessity to explain a multitude of experimental findings amidst experimental evidence in favor of distinctly non-Fermi liquid phases of fermionic matter. The most famous are certainly high temperature superconductors, cuprates and pnictides being the leading members of this heterogeneous group. The superconducting order at relatively high temperatures is almost the least important of the many unusual properties. A glance at the phase diagram of cuprates (Fig. 1.3) reveals how the doping of external charge carriers turns the system from the familiar normal metal, i.e. Fermi liquid phase into a non-Fermi liquid, universally known in condensed matter physics as strange metal, continuing on into the pseudogap. The pseudogap region is also mysterious, but thought to display some kind of long-range order. Dozens of exotic order parameters were proposed to explain this novel ground state: stripes, current loops, exotic spin ordering and others [118]. Many of them

⁶As a side remark, we refer the reader to [119, 72] for a possible geometric interpretation of the "fermionic constraint" which allows one to treat the problem of fermionicity by looking at certain global (topological) properties of the many-particle wave function and the path integral. The key result is the proof [13] that the signful path integral can be turned into ordinary bosonic path integral but with an additional constraint. For us, it is the morale and not the details which is important: it provides a construction which explicitly reduces fermion dynamics to boson dynamics with constraints. While AdS/CFT handles fermions in a somewhat different way, essentially trading the fermionic physics for curved space, it might be an indication that in the end all fermionicity can be bosonized by adding additional constraint structure to dynamics.

have some degree of experimental support [118]. We also do not understand how the advent of the strange metal is related to superconductivity itself.

A possible unifying point for these phenomena in unconventional superconductors and heavy fermions (as well as some other materials) is quantum criticality [95, 15], developed mainly by Sachdev. Its basic idea is that quantum fluctuations can mimic the effects of temperature on the order parameter of some ordered phase. The outcome is that the ordered phase becomes unstable and vanishes at a critical point at zero temperature. In place of temperature, the control parameter is typically some quantity which governs the competition between two ordered phases at $T = 0$, e.g. coupling strength or doping. The phase diagram of a system with a quantum critical point typically looks as in Fig. 1.3(A), quite similar to the phase diagram of real-world cuprates in Fig. 1.3(B) – above the critical point one has a characteristic quantum critical “cone”, the regime in which the quantum critical point influences the physics even at relatively high temperatures. This is an important difference with respect to finite temperature critical points: in the latter case, the scale invariance inherent to criticality is only felt in a narrow window around $T = T_c$, while quantum critical behavior can be detected even by measurements significantly above $T = 0$. Systems with a quantum critical point are mainly recognized for exhibiting remarkable scaling laws [111]. Of course, quantum criticality immediately brings associations on CFT and makes a great starting point for a holographic investigation. Notice the inverted epistemology of holography compared to conventional methods: normally, we would start from the Fermi liquid phase and try to build up interactions that drive it to the critical point. In AdS/CFT, we can start from the quantum critical point where the theory is very strongly coupled and completely encapsulated in the scaling relations, with particles being non-existent, and the challenge is to see how the system picks a ground state away from criticality. This is the essence of AdS/CMT: we know most in AdS/CFT precisely in the situation when we know least in conventional CMT.

This in turn makes CFT an important tool for description of such systems – scale invariance of the quantum critical phase is almost equivalent to conformal invariance. Therefore, if the universal ingredient in transitions from Fermi liquid non-Fermi liquid systems is quantum criticality, then the CFT and its gravity dual in AdS present a natural starting point.

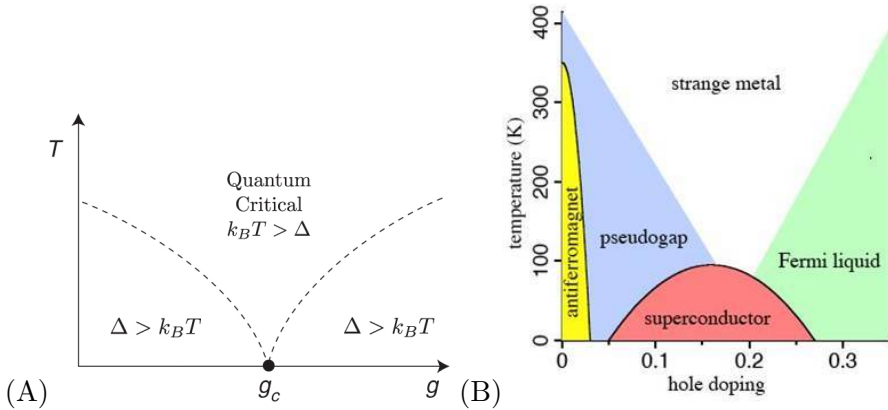


Figure 1.3: (A) Sketch of the phase diagram of cuprates. Normal metal phase turns into a non-Fermi liquid at critical doping, which presumably corresponds to a quantum critical point. The zoo of exotic orderings resides in the strange metal phase, which partially overlaps with the superconducting region. The properties of the strange metal remain mystifying, and might constitute a prime example of a stable non-Fermi liquid ground state. In (B) (adopted from [95]) we see schematically how the quantum critical point influences the physics at finite temperature: when the temperature energy scale $k_B T$ is larger than some characteristic energy scale Δ of the system, we are in the quantum critical "cone" where the physics is governed by the scaling laws imposed by the quantum critical point.

1.4 Outline

Our first goal will be to find the gravity dual for the Fermi surface which will be as general as possible and not hinge on existence of quasiparticles. The minimal ingredients we need are fermions, temperature and chemical potential. Our holographic dictionary translates this into a charged black hole plus the Dirac equation for the fermions. This is done in the next chapter. A Fermi surface should reveal itself in the spectrum of perturbations. We will study in detail the momentum and energy distributions of the spectral weight and conclude a great deal about the quantum critical fermions in this way. This chapter is adapted from [17] and includes the formalism for calculating the spectrum in a separate section (originally the Supplementary material of the paper). Chapter 4 [18] studies fermionic instabilities, giving dictionary entries for fermion density and Fermi liquid

itself, within a model we call black hole with Dirac hair. Chapter 5 [19] is the beginning of the study of the phase diagram of holographic fermions. In this chapter, we compare the Dirac hair model to the electron star model by Hartnoll et al [51] and show how one can interpolate between the two, corresponding to stable quasiparticles with different properties in field theory. In chapter 6 [83] we will study the actual phase diagram of holographic fermions by a full quantum-mechanical formulation of electron star and Dirac hair. and finally address the question – can we see novel phases from AdS/CFT? Chapter 7 sums up the conclusions.

