Lattices and Superstrings

Terry Gannon
Department of Mathematics and Statistics
McGill University, Montreal

23 January 1989

A Thesis submitted to the Faculty of
Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of Master of Science

Abstract

Modern particle physics and, in particular, superstring theory are reviewed. There follows a discussion of lattices, including self-duality, gluings and root lattices — concepts central to the subsequent material. Two main questions arising from recent work in superstrings (the self-duality of gluings; basis transformations in "compact lattices") are then solved. I also prove that every lattice is a sublattice of some cubic lattice Z^m, and I discuss the self-duality of the momentum lattice. In addition, scattered throughout the thesis are a number of smaller results.

La physique moderne des particules et, en particulier, la théorie des "superstrings" sont revues. Suit une discussion sur les réseaux, incluant l'auto-dualité, les collages et les réseaux de racines — concepts centraux au matiériel subséquent. Deux questions principaux sur des travaux récent en théorie des "superstrings" (l'auto-dualité des collages; les transformations de base dans les "réseaux compacts") sont ensuite résoudues. J'aussi prouve que chaque réseau est un sous-réseau de réseau cubique Z^m, et je discute l'auto-dualité du réseau moment. Nombre de plus petits résultes sont éparpillés dans tout la dissertation.

PREFACE

This thesis is concerned with superstrings, lattices, and some of the applications of lattice theory to the study of superstrings. Some of the following material is a survey of known results, some of it is original, and some is in between. In the next few paragraphs I hope to summarize what I have written, trying to separate what I have figured out from what I have borrowed.

In short, I first survey what is known in superstrings (Chapter 1) and lattice theory (Chapter 2); I then address two "big" problems (the self-duality of gluings in Chapter 3, and basis transformations of toroidal or compact lattices in §§1-3 of Chapter 4), two "medium" problems ("Every lattice is a sublattice of Z^m" in §5 of Chapter 2, and the self-duality of the momentum lattice in §4, Chapter 4), and a number of smaller results scattered throughout the last three chapters.

Chapter 1 is a discussion of modern particle physics and, in particular, the controversial theory of superstrings. In §1 I give a non-technical outline of what superstring theory is all about: what it attempts to do, what scientists like and dislike about it, and a brief sketch of its historical development. §2 is a slightly more detailed investigation into many of the key concepts in quantum field theory. There I address concepts, like anomalies and renormalization, which were referred to in §1, and also others, like symmetries and their representations, that will appear in the following, more detailed sections on superstrings. The material in these two sections come from a large number of sources, some of which are listed in the bibliography (see, for example, DB).

In §§3 and 4 I borrowed heavily from GSW and from my notes to a course[1] taught in Fall of 1988 by Dr. Jacques (which in turn was based on GSW). In §3 I begin the formal study of the string. Included is a discussion on why strings are studied over membranes or blobs. §4 addresses some more specialized topics, such as supersymmetry, the heterotic string, and modular invariance of the partition function.

The first four sections of Chapter 1 thus provide a general overview of superstrings. In the later chapters we will be concerned with specific aspects of the theory; the necessary background is provided by §5 of Chapter 1 (as

[1] 189-723A Representations of Lie Algebras with Applications to Physics

well as by §1 of Chapter 3 and §§1 and 4 in Chapter 4).

Chapter 2 provides us with a survey of the relevant topics in lattice theory. What GSW was to Chapter 1, CS is to Chapter 2.

§1 introduces lattices and some of the basic concepts. §2 discusses the important notion of self-duality. §3 discusses two important ways lattices can be "equal": congruence and similarity. §4 is concerned with the direct sums of lattices.

In §5 I ask the question: to what extent (if any) can we say each lattice can be embedded in some cubic lattice Z^m. I answer this, and with a series of counterexamples show that my answer is the strongest we can expect. This result also has a consequence for the factorization of symmetric matrices.

§6 describes an important family of lattices called root lattices. Gluing theory and the representations of Lie algebras are handled in §7, while in §8 we study the most famous lattice (the 24-dimensional Leech lattice) and a recursive process (called lamination) which generates it. (§§5 and 8 fall outside the main scope of this thesis).

Many of the results in Chapter 2 can be found in the literature, but I have been unable to locate proofs to a few of these. In those cases I have included my own proofs.

Chapter 3 is concerned with determining whether a given gluing is or is not self-dual. In §1 I discuss how this problem arises quite naturally from the study of superstrings. In the following two sections I give three methods for testing self-duality. The first two are entirely my own; the first accomplishes this by finding a basis for the gluing, while the second finds the determinant more directly. The second is particularly intriguing, and with it I establish with ease a few interesting results. The third method was proven by me, but was inspired by the work of Dr. Lam. Each of these methods have their own strengths and weaknesses. In the fourth and final section of Chapter 3 I describe an elegant graphical technique (due to Dr. Lam) for computing the determinants of certain matrices. It is very useful, for example, when using the first and third methods.

In §5 of Chapter 1 I described a certain parametrization of string theories. This parametrization unfortunately isn't one-to-one; the first three sections of Chapter 4 address how different sets of parameters can characterize the same string theory. This amounts to finding all the basis transformations in "toroidal" or "compact" lattices, which can be thought of as the quotients of "true" lattices. This work is all my own. In the final section of this thesis

I present a plausibility argument for the universally accepted but (I believe) unproven claim that the momentum lattice be self-dual. It seems likely to most people that modular invariance of the partition function implies its self-duality; I show that the modular invariance of a closely related function would indeed imply its self-duality.

In Chapters 3 and 4 I also establish a number of smaller results (e.g. Theorem 3.3.3) that I have been unable to find elsewhere.

I'd like to thank my advisor Dr. Lam, without whose patient help this could not have been written.

Tables and Figures

Table of Contents

1 PHYSICAL PRELIMINARIES

1.1 The Theory of Superstrings

A headline in PhysRev G, a joke publication produced by University of Toronto physics students, reads: "SUPERSTRINGS Solve Cadbury Secret!". The article continues, "In a stunning blow to the physics community and the Cadbury corporation, U of T researchers ... have succeeded in applying the complex and controversial theory of *superstrings* to explain the mystery of the Caramilk bar."

The supporters of superstrings include some of the biggest names in physics (Weinberg, Salam, ...), as do the critics (Feynman, Glashow, ...). Some say that the transition from point to string may be "no less profound than the transition from real numbers to complex numbers in mathematics."[1] Others, less convinced, argue:

> ...years of intense effort by dozens of the best and brightest have yielded not one verifiable prediction ... For the first time since the Dark Ages, we can see how our noble search may end, with faith replacing science, once again.[2]

Why all the fuss? What is it about superstrings that has the physics community so sharply divided between excitement and fear?

There are 3 fundamental forces, according to modern physics (gravitation, the strong nuclear force, and the electroweak force. Another force, a weak short-ranged Higgs force, is predicted but not yet observed). There are 37 elementary particles (6 leptons, 3x6 quarks, 8 gluons, and a photon, a Higgs scalar, a graviton, and a W and Z boson), and there are something like 18 fundamental parameters. What superstrings claims to be able to (ultimately) do is to explain all this mess simply and eloquently (the oft quoted word is "beautifully"). There really is only 1 fundamental object (the string), superstring theorists insist. It can only interact by joining with another, or by splitting in half. And there is only one parameter (the string tension T, or Newton's gravitational constant G). Everything else, they suggest, is mathematically forced.

[1] p. 1251, WAL
[2] p. 7, GG

That is what the advocates believe. The critics whistle a somewhat different tune, however, but they do so with no less conviction.

It all began in the late 60's as physicists were trying to come up with a reasonable theory of the strong nuclear force. The weak nuclear force was being successfully associated with the electromagnetic. Gravity succumbed to Einstein half a century earlier, but the strong force held fast.

When we speak of "force", we usually mean something that pushes or pulls — accelerates — the object it acts upon. By "force" physicists now mean something a little different. To them, a force is simply the exchange of particles. They affect particle scatterings, as would be expected, but they also are responsible for all decays.

The photon and the graviton are the exchange particles for electromagnetism and gravity, the dominant macroscopic forces. Z^0 and $W^{\pm}$ are exchanged in the weak force, causing among other things the instability of the neutron and the luminous glow of my clock. The strong force produces the heat and light of the sun, and, we now know, is mediated by gluons.

But back in the 60's things weren't quite so clear. Among other things there were just too many hadrons (strongly interacting particles), and some had ridiculously high spin. One of the more promising possibilities was the bosonic string of Yoichiro Nambu. It pictured quarks as being at the end of a very short ($\approx 10^{-13}$ cm), elastic string. This seemed a very natural way of handling a force that at close ranges was negligible, but at longer ranges became stronger and more confining. It explained the extra hadrons as being different modes of rotation and vibration of the string. It also had very good high-energy behavior, something which distinguished it from most of the competing theories.

Theories in particle physics seemed plagued by infinities. The mass of the electron could be calculated to be infinity, as could its charge, even though the measured values were considerably smaller. A procedure, developed by Feynman and others and called *renormalization*, allowed physicists to get finite, meaningful answers out of the theory. I'll discuss renormalization in more detail in the following section, but essentially it is a recipe for coming up with sensible answers: write down the (incorrect) equations predicted by the theory; fiddle with them in certain mathematically implausible ways, and out will come your (correct) answer. If you bother to check the intermediate quantities in your calculation (i.e. the ones you're fiddling with), you'll find infinities. But if the final answers will always be finite, the theory has

predictive power and is called renormalizable.

By 1974 quantum chromodynamics (QCD) had been constructed and was rapidly becoming established as the correct theory of the strong nuclear force. Interest in the bosonic string died down.

QED (the first successful quantum theory of electromagnetism), the electroweak theory, and QCD all have infinities, but all are renormalizable. The bosonic string, on the other hand, was completely finite. Physicists differed on just how reasonable renormalization was (Paul Dirac, for example, was one of its critics), but all of them recognized that a finite theory was certainly more desirable — all other things being equal, of course.

In spite of this success, the string model of the strong force had a few problems of its own. It predicted a tachyon (a particle whose mass squared is negative, and hence travels at speeds greater than light). It only made sense in 26 dimensions, while reality was clearly 4-dimensional (3 space, 1 time). It also demanded the existence of a spin 2 particle. Despite the great proliferation of hadrons, none had been found with spin 2. And finally, experimental scattering data unambiguously revealed nature's preference for QCD over the bosonic string. Today no one disputes that QCD more accurately describes the strong interaction.

A handful of physicists (including John Schwarz) continued to work in string theory, but no longer with the goal of explaining the strong force. To them, the previously embarrassing spin 2 particle was none other than the graviton. Theirs was a quantum theory of gravitation, and a whole lot more.

No physical theory is as convincing and impressive as Einstein's general relativity, where gravitation is reduced to geometry. But the theory has a flaw that guarantees it has only limited validity: it cannot handle the very small.

This flaw is shared by all attempts at a quantum theory of gravitation. In short, quantum gravity is nonrenormalizable — it's infinite, and those infinities can't be swept away. This is because the graviton has such high spin, something we can do nothing about. But Schwarz and others were suggesting that the problems were due as well to a fundamental prejudice all those theories shared, and we could do something about that.

In QED, in the electroweak theory, in QCD, in the faulty theories of quantum gravity — everywhere you looked you'd see point particles. Elementary particles (e.g. electrons and quarks) were supposed to be dimensionless points, spheres of radius zero. Schwarz suggested that this was the

source of the infinities.

Some people (e.g. Heisenberg and Yukawa) had tried alternatives, like rippling membranes or pulsating blobs, but these all violated locality or had some other crippling disorder. More to the point, they were symptomatic of physicists' conviction that a particle had to be a concrete object. In string theories the electron isn't an *oscillating string*; it is instead the *oscillation itself*, the *wave pattern* on the (infinitely thin, usually circular) string. And that has made all the difference.

The string theories considered today differ from the old bosonic model in a number of ways. They are *supersymmetric* (see §5) (hence the name *super*string), and thus include fermions (like the electron and quark) as well as bosons (like the graviton and the photon). They no longer have tachyons. They require only(!) 10 dimensions, instead of 26. They are 10^{20} times shorter, on the order of the Planck length ($\approx 10^{-33}$ cm), and are *closed* — i.e. circular — rather than *open*, with two endpoints. The various charges and quantum numbers are located at the endpoints of the open string, while they are spread uniformly along the closed string, being more a quality of motion. And, their advocates claim, they can unify *all* the forces and particles found in nature.

Today's theory of particle physics is called the Standard Model and is described in the next section. Experimentally speaking, it has been enormously successful: for example, the calculated and measured values for the electron's magnetic moment (in natural units) are respectively

$$1.00115965246 \pm 0.00000000020$$

$$1.00115965221 \pm 0.00000000003.$$

Nevertheless, physicists are unsatisfied with it for a number of theoretical reasons. First and foremost, there's the problem with gravity mentioned earlier. Also, they'd like to unify the forces, and in so doing reduce the number of elementary particles and undetermined parameters to more manageable levels.

The Grand Unified Theories (GUTs), which unify the strong and electroweak forces, are one way to do this. They make some interesting predictions, such as the decay of the proton, but don't solve the problem of gravity. Recently a lot of attention has been given to supergravity theories. Like superstrings they can handle gravity, are finite, and require many dimensions

(11 for the most promising theory), but unlike superstrings they have problems with chirality (an 11-dimensional theory can't violate parity), and the attention directed on them has since been shifted to superstrings.

The theory of superstrings is currently incomplete. Many important and encouraging results have been obtained, but much more work lies ahead. For example, John Schwarz is hopeful that by the turn of the century there will have been enough successes that its validity would be clear, but even so believes superstrings won't see their greatest hour until well into the 21st century. One of the problems is that some of the mathematics hasn't been developed yet. This has been cause for frustration for those anxious to see experimental proof of superstrings.

One of the most important properties of superstrings is their remarkable uniqueness: there are very few candidates that have any hope of being physically or mathematically acceptable. Gravitation is a forced characteristic of these theories. However, the uniqueness and simplicity of string theory — its "beauty" — seems hopelessly lost when you try to account for the extra dimensions of the theory: depending on how you count, the number of reasonable candidate string theories is either around 6 or in the thousands. Still, there is reason to believe that both these numbers can be reduced as deeper understanding of compactification is achieved.

The number of theories and the complexity of the math has made it very difficult to make accessible experimental predictions. The natural energy scale of the theory is Planck's mass ($\approx 10^{19}$ GeV) — you'd need a particle accelerator at least 10 light-years in length to get up to those ranges. One prediction that seems reasonable is the existence of another heavy photon (like Z^0), but no one is sure how heavy it would be. If a particle could be found weighing the same as a bacterium but with a charge a tiny fraction of the electrons, that would be a major victory for superstrings. Some models predict "shadow matter", matter that can interact with us only gravitationally; it would be invisible, and would pass right through us. However, only massive clusters of shadow matter could be directly observed, due to the weakness of gravity. But realistically there is little hope for the immediate contact between superstrings and experiment, and this has many physicists (like Feynman and Glashow) concerned.

Theoretical problems that the theory must explain include why the cosmological constant (a measure of the energy density of the vacuum, or the curvature of empty space) is measured to be so close to zero (it should be

zero as long as the theory stays supersymmetric, but for experimental reasons the theory can't), and why there are 3 identical (apart from mass) families of electrons and quarks (it seems to be related to the number of "holes" in the 6 extra dimensions). But the most obvious one is why reality seems 4-dimensional. One possibility is to imagine these extra dimensions as curled up (*compactified*) into tiny balls. These balls would have radii on the order of 10^{-33} cm, and so would be unobservable. The compactification would have occurred briefly after the Big Bang, or, as some suggest, may have triggered it.

There are problems, though, with straightforward compactification (see §4). An alternative is to not interpret the extra dimensions geometrically, but rather to treat them merely as interior degrees of freedom.

In superstrings, unlike general relativity, the formulas are coming first, long before the conceptual understanding of what the equations really represent. In other words, string theory is at this time little more than a set of computational rules, with no known underlying principles. This is disturbing many people. Says Edward Witten, one of the architects of the theory,

> But the fact that these things work, that these seemingly bizarre rules give ways of computing things in quantum gravity, giving sensible results and finite answers and leading in many different directions to all kinds of beautiful areas in mathematics, is a very deep mystery — probably one of the deepest which has ever been encountered in physics. It is unlikely that a proper understanding of this mystery will be found either soon or simply. But it will be worth the wait.[3]

Superstrings is a theory with much promise, certainly our greatest hope yet for a theory of everything. But whether it is an accurate theory of nature is a question we simply cannot answer yet. We could be living during one of the rare revolutions in physics, or we could be experiencing the much more common 'false start'. Only the theorists and experimenters in the next several years will be able to provide an answer.

If correct, superstrings will have accounted for, *in principle*, all of the effects in nature. It won't solve all mysteries, not by a long shot: our camping trips will still be ruined by inaccurate weather reports; the origin of life and

[3] p. 107, GHT

the nature of the mind will remain as unexplained as ever. The standard analogy is that it's like having learned the rules to chess — you're still a long way from being a grandmaster.

But that prospect doesn't excite me too much. What interests me is the search for the fundamental. It is only that which could have captivated men like Einstein. I don't mind if superstrings gets enshrined into established physics — not at all — but I for one hope it won't stay there.

The conclusion to a project done by me in a course earlier this year, written on the desirability to science of controversies like the fifth force, is perhaps even more appropriate here:

> May no physics paper, however desperately, ever end with:
>
> Amen! Amen! Amen! ,
>
> and may Science, that employer of humanity's finest, may Science never die.

1.2 Quantum Field Theory and the Standard Model

The two great pillars of modern physics are relativity (both special and general) and quantum mechanics. Relativity eloquently examines space and time and gravity; quantum mechanics speaks of the very small, and in so doing demolishes one by one some of our deepest and most cherished intuitions about what reality is really like. They were created independently early this century, and it was only in 1929 that their (partial) unification was accomplished, by Paul Dirac. Specifically, Dirac constructed a (special) relativistic theory of the electron. The complete unification of special relativity and quantum mechanics is now realized in the framework of quantum field theories, but general relativity stubbornly refuses to cooperate (unless we heed the victory shouts of the superstring enthusiasts).

The reason relativity and quantum mechanics are so difficult to bring together is that they are nearly incompatible. Common features can be abstracted from their very few acceptable unifications; hence fundamental concepts (e.g. spin and antiparticles) are derivable, rather than being ad hoc additions.

Quantum mechanics has largely been unchallenged and unchanged by the more modern theories. Its ideas concerning the fundamental indeterminacy of nature, or the superpositions of states, survive; the measurement

(pseudo-)problem is as much a (pseudo-)problem in the quantum mechanics of Schrodinger or Heisenberg as in the Standard Model or superstrings. String theorists, for example, claim that general relativity will arise asymptotically from superstring theories, just as Newtonian gravity arises in the limit from general relativity. But quantum mechanics, on the other hand, is inextricably interwoven into superstrings. It's the bedrock, it's the language, of modern physics.

More precisely, physicists have continually changed their minds about the dynamical behavior of elementary systems: first it was given by Schrodinger's equation, then Dirac's, then by the various models of quantum field theory culminating finally in the Standard Model. (Hence) they also have changed the Hilbert space of physical states. There have been some other changes (e.g. the idea of a particle in the early quantum theory differs from that in the quantum field theories, which in turn differs from superstring's suggestion). But the conceptual foundation upon which they have been built has always been quantum mechanics. The theory of superstrings is no exception.

Historically, field theory was introduced in the study of electromagnetism and gravitation to avoid the distasteful possibility of action-at-a-distance; the field would be the medium through which a disturbance would propagate outward with finite velocity. And in more modern physics our distaste for action-at-a-distance has (thanks to relativity) grown, so similar reasoning suggests fields be used in quantum theories as well. What makes fields a practical necessity is that in relativistic point theories the number of particles must be conserved. This is most definitely not the case in elementary particle physics (e.g. the decay of a neutron).

A field is simply an object (or set of objects) defined at each point in space-time. For example, the classical electromagnetic field consists of a (real) 4-vector $A_\mu(x)$ defined at each point x in space-time. A particle, finding itself immersed in the field, experiences a force proportional to the strength of the field at that point. What distinguishes a quantum field from a classical one is that the objects in quantum field theory are Hermitian operators, as opposed to numbers. The electromagnetic field becomes, for each space-time point x, a 4-vector $A_\mu(x)$ of Hermitian operators. These operators act on a Hilbert space of state vectors. The quanta of a field — i.e. the discrete energy states of the field oscillators — are identified with particles (e.g. the quanta of $A_\mu(x)$ are photons). This will be shown in a little more detail in the following section in the special case of superstrings.

Symmetries and their representations are important aspects of quantum field theories (e.g. the Standard Model), and we will now turn our attention to them. In order to account for experiment, a theory must be relativistic, and in order to be relativistic its symmetries must include the Poincaré group. This implies the following:

Let $u_\alpha(x)$ be a collection of fields in the theory, and consider any Poincaré transformation from one coordinate system x to another x', satisfying

$$x'_\mu = L_{\mu\nu}x^\nu + a_\mu$$

for some 4-vector a_μ (characterizing a translation) and a Lorentz transformation L. Physically this corresponds to choosing to describing an event using a different reference frame — perhaps one located in a different spot, in a different time-zone, in a laboratory rotated and moving uniformly relative to ours. Then there exists a matrix Λ (depending on L and a) such that under that Poincaré transformation $u_\alpha(x)$ gets mapped to $u'_\beta(x') = \Lambda_{\beta\alpha}u_\alpha(x)$. It is possible to show that $\Lambda(L,a)$ forms a linear representation of the Poincaré group. These representations are characterized by some non-negative $s \in \frac{1}{2}Z$ (s is called spin, though technically spin is measured in units of $\hbar$); if s is an integer the corresponding fields are called tensor fields (e.g. $A_\mu(x)$ is a vector field since the photon has spin 1), otherwise they are called spinors. For example, $A_\mu(x) \longrightarrow A'_\nu(x') = \frac{\partial x'_\nu}{\partial x_\mu}A_\mu(x)$.

To repeat this important point, the fields transform under the symmetries of the theory according to representations of the symmetry group. This applies not only to Poincaré symmetries, but to internal symmetries as well. Elementary particles correspond to irreducible representations of the Poincaré group. Kinematic quantities like momentum, mass, energy and (as we have seen) spin arise from this correspondence. Similarly, non-kinematic quantities like charge and lepton number arise likewise from other symmetries (though, unlike the Poincaré ones, they won't be symmetries of space-time).

It is a theorem in quantum field theory that the tensor fields will obey Bose statistics (e.g. will be gregarious), while the spinors obey Fermi statistics (and the Pauli exclusion principle). They are called bosons and fermions, respectively. Photons and gluons are examples of bosons. They are particles of force (e.g. two neighbouring electrons feel each other's charge by exchanging photons). Electrons and quarks are fermions (as are protons and neutrons) and are the particles of matter.

In quantum mechanics the observables are associated with Hermitian operators and also act on the states. The state vector (or "ket") stores all the physical aspects of a given system. Decomposing it into a sum of eigenstates of a given operator (e.g. momentum) tells you as much as is possible to know about a subsequent observation: the coefficient in front of an eigenstate yields the probability that the eigenvalue of that eigenstate will be the numerical result of the measurement. A simple example of a state is the vacuum $|\Omega\rangle$. It has zero total momentum and angular momentum (e.g. it is an eigenstate of the momentum operator, corresponding to eigenvalue 0). There are also operators giving the total number of particles in a given state (these operators as we shall see are intimately associated with the fields). Their eigenvalues are the natural numbers 0,1,2,.... $|\Omega\rangle$ is also an eigenstate of them, with eigenvalue 0. (This will be done more explicitly in the following section.)

There are creation and annihilation operators a and $\bar{a}$ (these are just the 'amplitudes' of the particle fields discussed earlier). They increase or decrease by 1 the number of (anti-)particles of a given type (e.g. "electron") and momentum. Applying these finitely often to the vacuum generates all physical state vectors (again, see the following section). From these creation/annihilation operators can be constructed the "number" operators discussed earlier.

What a quantum field theory must be able to calculate are transition amplitudes. Suppose we start at time t_0 with a state $|A\rangle$ and we want to know what the probability is that at time t_1 the system (if measured) will be in state $|B\rangle$. This is given by $|\langle B|A\rangle|^2$ ($\langle B|A\rangle$ is just the inner product the Hilbert space of states by definition possesses). For instance, $|A\rangle$ may denote an isolated neutron and $|B\rangle$ may consist of an electron, a proton, and an anti-electron neutrino. Then $|\langle B|A\rangle_t|^2$ will give the probability of the neutron having made the indicated decay by time t_1. Transition amplitudes are the means through which a quantum field theory confronts experiment. From them are calculated scattering cross-sections, decay rates, etc.

To each quantum field theory is assigned an object called a Lagrangian density $\mathcal{L}(x)$. It must be a function only of dynamical variables (i.e. the components $u_\alpha(x)$ of the fields, and their derivatives). It should be invariant under the Poincaré transformations so quantities such as energy and momentum will be constants of motion. It should be local (i.e. have no integrals in it), and it should be Hermitian (so that quantities such as energy and momentum will be real). The Lagrangian characterizes the dynamical be-

haviour of the theory, and can be used to calculate the transition amplitudes discussed above.

The Lagrangian also appears in classical mechanics, though in a slightly different way. The classical field obeys the action principle $\delta I = 0$, where I is the action

$$\int d^4x \mathcal{L}(x).$$

This leads to the well-known Euler-Lagrange equations of motion. The quantum field behaves differently. The important quantities that must be computed are the transition amplitudes, given by the Feynman path-integral. It is just the weighted sum over all possible (classical) paths going from the initial to the final states, the weight being just $\exp(iI/\hbar)$.

In general these calculations are far too complicated to yield exact answers, so what is used in practice is a perturbative expansion. A very elegant way of doing these perturbative calculations is due to Feynman, and involves Feynman diagrams. Basically what you do is imagine all possible ways of starting with $|A\rangle$ and ending with $|B\rangle$, apply Feynman's rules to $\mathcal{L}$ to calculate the transition amplitudes of each of these possibilities, and then take the (infinite) sum. Fortunately the more complicated possibilities tend to contribute little to the sum (at least in theories like QED, unlike QCD), so only the simplest few need be considered. In particular, the number of "loops" in the diagram corresponds to the order of that perturbation term.

In all cases of physical interest $\mathcal{L}$ is not only invariant under Poincaré symmetries, but also under so-called internal symmetries. The fields behave under these symmetries analogously to how they behaved under the Poincaré ones. These symmetries generally differ from the Poincaré ones in two ways: they often are gauge, as opposed to global, symmetries; and they usually don't affect space-time points. The Poincaré transformations (L, a) were the same at each space-time point x — i.e. L and a were constants. But for gauge symmetries the transformation is allowed to vary from point to point. The difference is a profound one (e.g. the difference between special and general relativity, or between supersymmetry and supergravity). Not all internal symmetries need be gauge ones. Some (e.g. parity) may not even be continuous.

Today's official quantum field theory is called the Standard Model. It is given by a symmetry group (namely, $SU_c(3) \times SU_L(2) \times U_y(1)$) and a set of elementary particles (6 leptons, 6 quarks, 8 gluons, the photon and $W^{\pm}$

and Z^0 bosons, and a Higgs scalar). $SU_c(3)$ characterizes the strong nuclear force, while $SU_L(2) \times U_y(1)$ is the gauge group of the unified electroweak force. The Standard Model also tells you how these particles behave under the symmetries. For example, the left handed particles $\binom{\nu_e}{e}$ transform as an $SU_c(3)$ singlet, an $SU_L(2)$ doublet, and have eigenvalue $-\frac{1}{2}$ for the generator of $U_y(1)$. In other words, they are unaffected by $SU_c(3)$ (which means they don't feel the strong force), but they do take part in electroweak interactions, and in fact have weak hypercharge $-\frac{1}{2}$ and their transformation matrices Λ characterizing their behavior under $SU_L(2)$ form a 2-dimensional irreducible representation of $SU_L(2)$.

To be able to compare the theory with experimental results we need a Lagrangian $\mathcal{L}$. $\mathcal{L}$ will be built out of the particles given above, and must be invariant under the symmetry group. Given the gauge group and the list of particles (and their transformation properties), if we are to have a renormalizable theory, the specification of something like 17 parameters (e.g. the mass of the electron) are required to determine $\mathcal{L}$ uniquely. Most of these have to do with the Higgs particle, which was introduced in order to give the relevant particles a mass — the Higgs is used to break the electroweak symmetry (this process will be explained shortly). This gives (large) mass to the vector bosons $W^{\pm}$ and Z^0, hence explaining why the weak force should be so weak and short-ranged when electromagnetism is not. The resulting symmetry group is $U_{em}(1)$, the gauge group for electromagnetism. This symmetry is exact, so the photon is massless (thus travelling at the speed of light - - a fortunate thing indeed!), and electric charge is conserved.

The symmetry $SU_c(3)$ remains unbroken, so gluons, the exchange particles of the strong force, are massless. The theory of the strong force is called QCD; the Standard Model doesn't attempt to unify the strong force to the electroweak ones (this is one of the reasons why the Standard Model isn't considered a final theory). The strong force has three types of charge, called *colour*. Leptons are colourless, but gluons are not, so QCD (like general relativity but unlike QED) is nonlinear.

Particles connected by a symmetry have the same mass. This is one reason to break a given symmetry. There are two ways to do this. One, called dynamical symmetry breaking, involves adding a small term to the Lagrangian which isn't invariant under the given symmetry. In other words, the symmetry wasn't truly exact.

The more important way for our purposes is called spontaneous symmetry

breaking. It occurs if the vacuum of the theory isn't invariant under the symmetry. The original (false) vacuum was invariant, but it wasn't a state of lowest energy. In effect the theory exchanges some symmetry for stability. The symmetry is still there, but is very well hidden. It can be re-established (in effect) if the energy is high enough. For instance, at about 100 GeV $SU(2) \times U(1)$ becomes unbroken and the electroweak forces blend together, while the unification energy for the GUTs is about 10^{15} GeV, and that for superstrings is around 10^{19} GeV.

There are two types of mathematical problems that plague quantum field theories. One is a problem of infinities, and the other is a problem associated with symmetries.

Quantum field theories have enormous difficulty staying finite. Some of these infinities can be mathematically removed by processes called renormalization; those that can't are called non-renormalizable. The philosophy behind renormalization is that although the results of the calculation of a physically measurable quantity must be finite, infinities may arise in intermediate steps. The Standard Model is renormalizable: sensible answers can be squeezed out. Quantum gravity, unfortunately, is not (this is related to the fact that the graviton has such a high spin).

The justification for renormalization is that our theory is not of unlimited validity; there's a "cutoff" distance corresponding to the smallest distance (or highest energy) to which our theory can be taken seriously. For classical theories (like general relativity) this consideration isn't important, but in quantum calculations even the very small distances (relative to the cutoff), corresponding to virtual processes of very high energies, contribute non-negligibly. A renormalizable theory is one that requires only finitely many parameters to handle the observable effects of these infinitesimal distances. It's a theory in which the energy level of the hidden, underlying theory of physics is much higher than that of the process being studied (e.g. the old theory of the weak force wasn't renormalizable because the mass of the W and Z bosons, though large, aren't large enough).

Symmetries are important for a number of reasons, but one of the most important is because they imply conservation laws. Electric charge is conserved because the Standard Model (Lagrangian and vacuum) is invariant under $U_{em}(1)$. An anomaly is a symmetry of the classical theory which is lost as you pass to the quantum theory (i.e. as you "quantize" the classical system) — in other words, the quantum corrections don't respect the desired

symmetry. There are many types of anomalies, and the presence of any of them is a serious problem for the theory. It turns out that non-chiral theories (i.e. theories invariant under parity) don't have problems with anomalies, but chiral theories are plagued by them. Since experiment demands we have a chiral theory, anomalies can present major dificulties. It turns out that, surprisingly, the Standard Model is anomaly-free.

Green and Schwarz' 1984 calculation on anomaly cancellation in string theories was the direct cause of the recent explosion of interest in superstrings.

Perhaps a word should be mentioned about the units used in particle physics. Usually the "natural units" are used: $c = 1 = \hbar$. I will adopt this convention for the most part. This means that energy and mass and inverse distance and inverse time will all have the same units. Of course, when physicists need to get actual numbers out of a calculation, to compare its predictions with experiment, they must insert (using dimensional analysis, say) c and $\hbar$ into the relevant parts of the equations, but this isn't difficult.

1.3 An Introduction to String Theory

In this section we'll begin a more mathematical study of string theory. We'll lead into it by first considering a more familiar quantum field theory.

Consider first a massless classical point particle. It traces out a world line in space-time. Its action is given by

$$I = \int \frac{dx_\mu}{d\tau}\frac{dx^\mu}{d\tau}e^{-1}(\tau)d\tau$$

where τ is a parameter running along its worldline. $x(\tau)$ is the position of the particle in space-time when the parameter equals τ. e is introduced to make I independent of the specific parametrization chosen — it's a kind of 1-dimensional metric along the world line. These reparametrizations are a gauge symmetry of this theory; we will now indicate a standard way to handle theories with gauge symmetries.

Fix the gauge: take $e = 1$. Then the Euler-Lagrange equations (which are derived from the classical action principle $\delta I/\delta x^\mu = 0$) lead us immediately to the equation

$$\frac{d^2x_\mu}{d\tau^2} = 0,$$

for each $\mu = 0, 1, \ldots, D-1$ (let D be the dimension of space-time; we won't commit ourselves yet to $D = 4$, but we will commit ourselves to a Minkowski metric). That is, this (free) particle will travel in a straight line (i.e. along a geodesic) in space-time.

But we can do more. We haven't exploited yet the gauge symmetry. I is invariant under reparametrizations, so we also have $\delta I / \delta e^{-1} = 0$. This yields

$$\frac{dx_\mu}{d\tau}\frac{dx^\mu}{d\tau} = 0,$$

so we get that the massless classical particle travels along null geodesics.

Now let's consider the transition to a quantum theory. x^μ and the canonical momenta $p^\mu = \frac{dx^\mu}{d\tau}$ now become operators: $\hat{x}^\mu$ is multiplication by x^μ, and $\hat{p}^\mu$ is the differential operator $-i\frac{\partial}{\partial x_\mu}$. An additional change is that the classical Poisson bracket becomes a commutator. The particle is now represented by a field $\phi(x^\mu)$.

The gauge symmetry now yields the expression

$$\eta^{\mu\nu}\frac{\partial^2\phi}{\partial x^\mu \partial x^\nu} = 0,$$

known as the Klein-Gordan equation. However, as was mentioned in the last section, quantum mechanics doesn't obey the action principle, so there is no analogue to the geodesic equation derived above (to a certain extent, though, the quantum particle's world line can be pictured as a fuzzy cloud centered about the null geodesic).

Now let's consider the classical string. The obvious question is what to choose for its action. Note first that there is a simple geometric interpretation of the action for a classical point particle: it's just the length of the world line (this is more clearly seen by investigating a massive particle).

We'll be open-minded, at least for a while, and discuss the theory for all higher-dimensional generalizations of a (0-dimensional) point particle. Let n denote the dimension of the fundamental object of the theory (so point particles are $n = 0$, strings are $n = 1$, and membranes are $n = 2$). Again, the dimension of the background space-time will be D, and its points will be called x^μ, $\mu = 0, 1, \ldots, D-1$. The trajectories of these objects will be an $n+1$-dimensional submanifold in this background space, and σ^α for $\alpha = 0, 1, \ldots, n$ will label its points. The (Minkowski) metric of space-time will again be

denoted by $\eta_{\mu\nu}(x)$, while $h_{\alpha\beta}(\sigma)$ will be the induced (Minkowski) metric on the 'world manifold'. Then, proceeding by analogy with the case given above, we are led to consider for the classical action the $n+1$-dimensional 'volume' of the world manifold:

$$I = -\frac{T}{2}\int d^{n+1}\sigma\sqrt{h}h^{\alpha\beta}(\sigma)\eta_{\mu\nu}(x)\partial_\alpha x^\mu \partial_\beta x^\nu,$$

where h is the absolute value of the determinant of $h_{\alpha\beta}$. x^μ is a function of σ — it is the position in space-time of the given point on the world manifold.

Now, the world manifold must not be observable, for it is merely an auxiliary object. Thus neither its coordinates σ^α nor its curvatures (which are computable from $h_{\alpha\beta}$) should survive the gauge symmetries of the theory. The presence of the metric h guarantees that I will be independent of any particular parametrization σ of the world manifold. This is good. But we must also be able to choose a gauge (as we did for the point particle) in which the metric itself is eliminated. h is symmetric, so it has $\frac{1}{2}(n+1)(n+2)$ independent components. But there are only $n+1$ independent gauge transformations (one for each parameter σ^α), so for $n > 0$ we can't quite eliminate the metric h with these reparametrizations alone. There is an additional symmetry for $n = 1$ (a Weyl rescaling) which completes the process for $n = 1$, but the higher dimensional theories can't be salvaged. And $n = 0$ (corresponding to the point particle) has renormalization problems with gravity. Thus we are led to consider $n = 1$, i.e. string theory.

Of course this 'no-go' 'theorem' doesn't irrefutably rule out higher dimensional theories than strings, but it does present a challenge to those who wish to construct non-string theories. In any event, from this point on only $n = 1$ will be considered. In addition, we will assume (as we do throughout this thesis) that the strings are closed (i.e. topologically a circle), with period π (the reason π, and not 2π, is usually chosen is (I believe) because historically the first string theories involved the open string, for which a period of π is quite natural, and the convention stuck). Unlike §5, for example, I will consider here only periodic boundary conditions.

Use the gauge symmetries then to map $h_{\alpha\beta}$ to the constant Minkowski metric (this is gauge fixing). We get as an equation of motion (varying I with respect to x^μ):

$$(\partial_0^2 - \partial_1^2)x^\mu(\sigma) = 0.$$

This can be trivially solved in the light-cone coordinates ($\sigma^{\pm} = \sigma^0 \pm \sigma^1$):

$$x^\mu(\sigma) = x^\mu_R(\sigma^-) + x^\mu_L(\sigma^+)$$

$$x^\mu_R(\sigma^-) = \frac{1}{2}x^\mu + p^\mu\sigma^- + i\sum_{n\neq 0}\frac{1}{n}\alpha^\mu_n e^{-in\sigma^-}$$

$$x^\mu_L(\sigma^+) = \frac{1}{2}x^\mu + p^\mu\sigma^+ + i\sum_{n\neq 0}\frac{1}{n}\bar{\alpha}^\mu_n e^{-in\sigma^+}.$$

Here α and $\bar{\alpha}$ are arbitrary complex numbers and have nothing to do with each other nor with the indices given earlier. The subscripts R and L denote right- and left-moving modes. x^μ and p^μ are the position and momentum of the centre of the string.

The gauge constraint is that the energy-momentum tensor $T_{\alpha\beta}$ must vanish, which leads to the equations

$$L_m \equiv \frac{1}{2}\sum_{n=-\infty}^{\infty}\alpha^\mu_{m-n}\alpha^\nu_n\eta_{\mu\nu} = 0 \text{ and}$$

$$\bar{L}_m \equiv \frac{1}{2}\sum_{n=-\infty}^{\infty}\bar{\alpha}^\mu_{m-n}\bar{\alpha}^\nu_n\eta_{\mu\nu} = 0.$$

T can be interpreted as the tension of the string; usually $T = 1$ or $T = \frac{1}{\pi}$ is chosen to simplify the notation (we will take $T = 1$ here). It is relevant only to the quantum theory, which we will now turn to.

The amplitudes α^μ_n and $\bar{\alpha}^\mu_n$ become operators, called oscillators. In order that $x^\mu(\sigma)$ be Hermitian, we must have

$$\alpha^{\mu\dagger}_n = \alpha^\mu_{-n}\ ,\ \bar{\alpha}^{\mu\dagger}_n = \bar{\alpha}^\mu_{-n}.$$

These satisfy commutation relations derivable from the Poisson brackets of the classical amplitudes.

These operators act on a space of states, an infinite-dimensional indefinite vector space called a Fock space $\mathcal{F}$. It can be defined as follows.

Assume the existence of *ground states*: i.e. vectors $|0;p\rangle$ (they have D components) satisfying

$$\alpha^\mu_n|0;p\rangle = \bar{\alpha}^\mu_n|0;p\rangle = 0 \text{ for } n > 0, \text{ and}$$

$$\alpha_0^\mu|0;p\rangle = \bar{\alpha}_0^\mu|0;p\rangle = p^\mu|0;p\rangle;$$

i.e. they're annihilated by the positive oscillators and are eigenstates of the $n = 0$ oscillators.

Then $\mathcal{F}$ is spanned by states of the form

$$\alpha_{n_1}^{\mu_1}\cdots\alpha_{n_k}^{\mu_k}\bar{\alpha}_{m_1}^{\nu_1}\cdots\bar{\alpha}_{m_\ell}^{\nu_\ell}|0;p\rangle.$$

Not all these states are *physical*; to be physical they must satisfy the gauge constraint. That is, they must be annihilated by L_m and $\bar{L}_n$ for $m,n \geq 0$ (L_m and $\bar{L}_n$ are defined as in the classical case, although for $m = n = 0$ this definition is slightly ambiguous).

It turns out that $L_0 - \bar{L}_0$ generates translations in σ^1, and hence can be interpreted as a sort of momentum operator for the string. $L_0 + \bar{L}_0$ generates translations in σ^0, and so can be interpreted as the Hamiltonian, or energy operator (strictly speaking we must add a multiple of the identity to get the Hamiltonian).

$$N = \sum_{n=1}^{\infty}\alpha_{-n}^\mu\alpha_n^\nu\eta_{\mu\nu} \text{ and } \bar{N} = \sum_{n=1}^{\infty}\bar{\alpha}_{-n}^\mu\bar{\alpha}_n^\nu\eta_{\mu\nu}$$

are called the number operators; for example, when $n_1,\ldots,n_k < 0$ a straightforward calculation gives

$$N(\alpha_{n_1}^{\mu_1}\cdots\alpha_{n_k}^{\mu_k})|0;p\rangle = (-\sum_{j=1}^{k}n_j)(\alpha_{n_1}^{\mu_1}\cdots\alpha_{n_k}^{\mu_k})|0;p\rangle.$$

It turns out that not quite all of our gauge freedom was used up by fixing $h^{\alpha\beta}$ =diag(-1,+1). The remaining freedom can be removed (there are alternatives) by a noncovariant gauge fixing called the light-cone gauge, which involves singling out the coordinates x^0 and x^{D-1}. The net effect is that the only independent oscillators are the transverse ones α_n^i and $\bar{\alpha}_n^j$ for $i,j = 1,\ldots,D-2$ — all others can be determined from these (or are zero) using the gauge constraints. This means that we should apply only the transverse oscillators to the ground states to get physical states.

Ghosts are physical states $|phys\rangle$ with negative norm : i.e. $\langle phys|phys\rangle < 0$. This is a very undesirable situation because of our wish to interpret these inner products as probability amplitudes. It turns out that in the light-cone gauge there are no ghosts. However, the gauge fixing broke Lorentz

invariance. You must make sure the resulting theory is Lorentz invariant. It turns out that Lorentz invariance holds only in $D = 26$ dimensions (i.e. in the other dimensions we get an anomaly).

Incidently, the string theory considered here is the free bosonic one. It is ***bosonic*** because of our choice of action I, and in particular our implicit assumption that the metric $h^{\alpha\beta}$ suffices to describe the geometry of the world sheet. More promising strings (the bosonic string has tachyons and no fermions) can be treated more or less similarly. It is *free* because we ignored interactions.

One of the appeals of string theory is the simplicity of its interactions: two strings may join into one; one string may split into two (and of course any combination of these may also occur). The resulting world sheet (speaking classically, for simplicity) of the scattering of two strings, say, may be two non-intersecting cylinders, or it may look like an 'H', or like a ladder with two rungs, etc. Quantum mechanically, we must take the weighted sum of these possibilitites. The result is a perturbation series. Which 'weight' to use can be determined from the Lagrangian (or the action) using Feynman's rules:

1. To each external tube associate an operator called the "vertex operator".

2. To each internal tube associate an operator called a "propagator".

3. To each loop take the trace of the corresponding product of operators in the space of states.

4. Integrate over the momenta in the loops.

The vertex operators and propagators can be calculated from the Lagrangian (for example, the propagator can be thought of as the functional inverse of the hamiltonian). We will see an important example of this in the following section.

1.4 Miscellaneous Topics in String Theory

In quantum mechanics it is possible to start with nothing and end with nothing, but to do so in a very complicated way. It does this through the so-called

virtual processes. In string theory, for example, what could happen is that a virtual string spontaneously appears, splits, rejoins, and then disappears.

These processes contribute to what is known in quantum field theory as the "vacuum-to-vacuum amplitude". Its first order term looks like a torus, and is called the partition function. The torus can be characterized by a complex number τ called the modular parameter, and so the partition function will be a function $Z(\tau)$, sometimes written as $Z(\tau, \bar{\tau})$, where $\bar{\tau}$ is the complex conjugate of τ but is treated as an independent variable (more accurately, the first term of the vacuum-to-vacuum amplitude is actually the sum over all tori: $\int d\tau Z(\tau)$ or $\int d\tau d\bar{\tau} Z(\tau, \bar{\tau})$, where the integrals are over the "fundamental domain").

Feynman's rules help us to arrive at (using the Dedekind eta function $\eta(\tau)$)

$$Z(\tau = x + iy) = \int d^{D-2}p \mathrm{Tr}\ \{e^{-yH} e^{ixP}\}$$

$$= |\eta(\tau)|^{-2(D-2)} (2\mathrm{Im}\ \tau)^{-(D-2)/2} \exp(-\frac{\pi}{6} \mathrm{Im}\ \tau (D - 26)).$$

Several different τ correspond to the same torus; Z should not be influenced by which of these equivalent parameters we choose. The group of transformations which map each τ to a τ' characterizing the same torus can be easily found, and is called the modular group. Thus $Z(\tau)$ must be invariant under the modular group — this is how the powerful constraint of modular invariance enters into string theory. It turns out that (for this bosonic string) modular invariance also demands $D = 26$. Many other consequences of modular invariance will be seen in §5, as well as in Chapters 3 and 4. Modular invariance guarantees the vacuum-to-vacuum amplitude (to first order) is finite. Some plausibility arguments exist which seem to show that modular invariance also guarantees that strings are free of all divergences, so that string theory is completely finite.

The comments thus far have concentrated on the bosonic string. In 1971 Ramond, Neveu and Schwarz found a fermionic string that was later discovered to have built into it a previously unknown symmetry called supersymmetry.

The symmetry SU(3) of the Standard Model links up the quarks with each other, and the gluons with each other. However, supersymmetry is the only symmetry that can mix bosons and fermions, and thus is our only hope to unify all the particles found in nature. It can do this because it has a

fermionic generator Q which changes the spin of particles by $\frac{1}{2}$. (See FRE for a complete introduction to supersymmetry.) Locally supersymmetric theories (called supergravity) automatically include general relativity and hence are serious candidates for a theory of quantum gravity.

There is no experimental evidence yet that nature is supersymmetric (in fact if it is, the supersymmetry must be badly broken), but many physicists are nevertheless convinced that supersymmetry is just too beautiful and promising not to somehow play an important role in reality. It was thought for a while that it might enter through supergravity, but there are now reasons for doubting this (the theory may not be finite, and its most promising versions aren't chiral). Today the best hope for supersymmetry seems to be superstrings.

There are several classes of supersymmetric strings — i.e. superstrings (the bosonic string has no fermions and so has no hope to be supersymmetric). Type I superstrings are both open and closed, while type II strings are only closed. The former seems to hold some promise as a possible theory of physics. The latter has difficulty either with chirality or with supporting an adequate gauge group.

But the most promising superstring today, first introduced in GHMR, is called the heterotic string (from the Greek word "heterosis", meaning the increased vigour displayed by crossbred plants or animals). It is a closed string, and hence its right- and left-moving modes are independent (there are no endpoints to reflect its wave). It is a hybrid of the old bosonic string and the type II string: its left-movers are bosonic and its right-movers are type II. This means that its left-movers live in 26 dimensions while its right-movers are only in 10, but this is rectified by making 8 of these transverse and 16 of them internal (2 are eliminated by the light-cone gauge). Only the right-movers are supersymmetric. The low-energy limit of the theory is $D = 10$, $N = 1$ supergravity (D is the dimension of space-time, N is the number of fermionic generators Q and the number of spin $\frac{3}{2}$ supersymmetric partners of the graviton, called gravitinos) coupled to the gauge group $\mathrm{Spin}(32)/Z_2$ or $E_8 \times E_8$. These gauge groups, and the way to rectify the difference of 16 dimensions are closely related to each other and to the 16-dimensional even self-dual lattices, and is one of the main ways lattices enter into string theory. The heterotic string is anomaly-free, free of ghosts and tachyons, and there is reason to believe it's entirely finite. Its lowest mass states (and there are many of them!) are all massless (which is good, since otherwise their masses

would be on the order of the Planck mass, $\approx 10^{19}$ GeV).

The most promising of the heterotic strings is the $E_8 \times E_8$ one. It seems to have the best hope of predicting the observed particles. It has been speculated that one of these E_8's might give rise to all of the observable particles, while the other would give rise to another type of matter (called shadow matter) which can interact with our matter only gravitationally.

The discovery of the heterotic string followed the explosion of interest in strings brought on in 1984 when Green and Schwarz discussed anomaly cancellation in type I strings (the only anomaly-free theories known until then were the unpromising type II strings). For example, they showed that cancellation of a certain gravitational anomaly could occur only when the gauge group was of dimension 496. Since E_8 and D_n=SO($2n$) are of dimensions 248 and $n(2n-1)$, respectively, the gauge groups given above avoid that anomaly. Anomaly cancellation enormously restricts the physically allowable theories and takes us a giant step forward to the dream of being able to derive a unique "Theory of Everything".

The heterotic string, like most of the superstrings being currently studied, is 10-dimensional (see §8, Chapter 2 for an example of a 26-dimensional theory). Yet nature appears to be only 4-dimensional. The most obvious way to explain the discrepancy is to *compactify* the extra 6 dimensions — to make them so small (e.g. on the order of the Planck length of 10^{-33} cm) that we have no hope of observing them. The problem with this approach seems to be that the $N = 1$ supersymmetry becomes $N = 4$ in the low energy limit, which is non-chiral and must be discarded. It turns out that the problem is that the tori considered here are flat. This has led to the consideration of Calabi-Yau manifolds, and of orbifolds, which are flat everywhere except at a number of singularities. An orbifold is the 'quotient' of a lattice with some subgroup of its automorphism group — it is mentioned here only for completeness, and won't be discussed again.

1.5 A Summary of Recent Work

As should be fairly obvious by now, my thesis is concerned with aspects of superstring theory — more precisely, my work in Chapters 3 and 4 is motivated by questions arising from the recent work in superstrings done by Lam (see LAM1-3). A similar approach was taken by the "Cornell group" (see, for example, KLT; the isomorphism between the two approaches was

made explicit in LAM1). In this section I will give a brief outline of some of these results, so that the work in Chapters 3 and 4 can be more properly put into perspective. See also §1 of Chapter 3 and §§1 and 4 in Chapter 4 for the specific physical questions I will address.

We are interested in the type II and heterotic string theories (these closed strings allow the cancellation of anomalies; an alternative, the bosonic string, is plagued with tachyons and so is a doubtful candidate for describing nature), and in how the various constraints (like modular invariance) restrict the physically allowable theories.

Let $X_\pm$ and $\Psi_\pm$ be boson and fermion fields, respectively. They are functions of (σ^0, σ^1) (in fact, of $\sigma^\pm = \sigma^0 \pm \sigma^1$), where $\sigma^0 \equiv t$ is time, and where $\sigma^1 \equiv \sigma$ is a parameter that runs along the string. Now, our strings are all closed, so σ is a periodic coordinate, say with period π. How X and Ψ behave as we wrap around the string — i.e. when we replace σ with $\sigma + \pi$ — constitutes their boundary conditions. As they aren't themselves directly observable, they don't have to be periodic.

The conformal currents (which generate the conformal transformations) look something like

$$T(\sigma, t) = -\frac{1}{2}\partial X^\mu \cdot \partial X_\mu - \frac{1}{2}\partial \Psi^\lambda \cdot \psi_\lambda.$$

They are physical, so must be periodic (i.e. $T(\sigma, t) = T(\sigma + \pi, t)$), which suggests the boundary conditions

$$X(\sigma + \pi, t) = \exp(-2i\pi\hat{w})X(\sigma, t) \text{ or } X(\sigma + \pi, t) = X(\sigma, t) + c,$$

$$\text{and } \Psi(\sigma + \pi, t) = \exp(-2i\pi w)\Psi(\sigma, t),$$

for constants c, w and $\hat{w}$. These are called the twist (by phases w and $\hat{w}$) and shift (by c) boundary conditions.

The superconformal current is a fermionic quantity, so it may be either periodic or antiperiodic. Without going into details, this condition relates some of the w to $\hat{w}$, and relates the other w's by the so-called triplet constraint (at least for $D = 4$). (In the heterotic string, unlike the type II one, only the right-hand side is supersymmetric, so only it has a superconformal current. Both sides are conformally invariant, though.)

There are two ways Ψ can be related to X. One way is via supersymmetry, and was discussed in the previous section. The other way is through

bosonization or its inverse, fermionization, which are given by the formula

$$\Psi =: \exp(-2iX) :$$

where the colons denote the normal ordered product, and tells you how to interpret products of fields. This relates the bosons satisfying the shifted boundary conditions to the fermions with the twisted boundary conditions. Fermionizing the former allows us to consider only the phase boundary conditions.

One thing that modular invariance tells us is that these phases must form an abelian group G (using addition modulo 1). We will assume them all to be rational numbers, so that G is also finite.

Physical states, unlike the fields, must be periodic. Given an arbitrary (not necessarily physical) state, the GSO projection allows one to project out the physical, i.e. periodic, component.

Modular invariance requires that there must be several types of boundary conditions (i.e. several different phases w). To each possible boundary condition (i.e. to each element $h \in$ G) there is associated a different Hilbert space of solutions (called a "sector") satisfying that boundary condition. Also, the modular invariance of the partition function demands that the phases w satisfy a number of relations, called (L), (Q), (O) and (N). See Chapter 4 for a discussion of the first three of these; together with anomaly cancellation, (N) puts a number of constraints on the numbers of right and left moving boson and fermion fields, and on the space-time dimensions of the acceptable theories.

Given a set of twist parameters w satisfying the above relations, a theory is specified by choosing various vacuum parameters F_i and m_{ij} (these specify the GSO projections permissible by modular invariance). From these we can calculate the spectrum and the symmetry group of the theory.

Space-time has a preferred position in quantum field theory. But here everything is an operator of the world sheet — including space-time. The Lorentz group is on an equal footing with all other symmetries of the theory. The spin-statistics theorem, giving the correct relationship between spin and statistics (see §2), is automatic in quantum field theory, but in string theory its validity isn't guaranteed: it's imposed, and not derived. Insisting upon it fixes the vacuum fermionic phases F_i. The result is that there are only finitely many physically acceptable theories allowed, given the twist group

G. (Presumably other demands not yet imposed will reduce the number even further). In §§1-3 of Chapter 4 we discuss how different sets of twist parameters can yield the same theories.

These results all followed from the fermionization of the bosons with shifted boundary conditions. If we instead bosonize all the fermions, we get equivalent but different conclusions. The main result is that the momenta of the bosons lie on a shifted lattice $\Lambda + t$. Λ turns out to be self-dual. In §1 of Chapter 3 and §4 of Chapter 4 I provide a more detailed account of this approach.

The bosonization approach is probably better because it allows you to handle much more easily the bosons with twisted boundary conditions -- these bosons in fact lead quite naturally to the study of orbifolds, which were considered briefly in the previous section. Also, it treats the whole subject in a much more uniform fashion: you don't have to break everything down into sectors. For example, the GSO projection is basically built right into the formulation, and the $E_8 \times E_8$ gauge group arises much more blatantly in the bosonization approach. On the negative side, the fermionization process presents some complications, as two real fermions with incompatible boundary conditions cannot always be bosonized (it takes two real fermions, or one complex one, to make a boson).

This is hardly intended to be an exhaustive survey of this recent work. The interested reader should consult the papers mentioned earlier for the details.

2 MATHEMATICAL PRELIMINARIES

2.1 Introduction to Lattices

Occasionally in mathematics a term is encountered that by itself represents several fundamentally different structures. The classic example of this is "field": to most mathematicians this refers to a certain algebraic structure analogous to the rational or real numbers. Mathematical physicists usually mean by this a vector or tensor field, as in "quantum field theory" or the expression "gravitational field". And apparently in set theory it is used to denote the union of the domain and range of a function or relation.

Another example of a mathematical homonym is "lattice". A lattice to most modern mathematicians involves two binary operations on a partially ordered set. This structure has applications in almost every field of mathematics (pardon the expression). In theoretical physics it can be found in the study of the foundations of quantum mechanics, and in quantum logic. This algebraic structure has absolutely nothing to do with the type of lattice concerned with here.

Definition 2.1.1 *A* lattice *is a finitely generated free Z-module, on which is defined a bilinear form.*

Examples of 2-dimensional lattices are given in Figures 1 and 4. (Another type of lattice, a toroidal or compact one, will be studied in Chapter 4 — it can be thought of as the quotient of two "true" lattices). We will use the symbol Λ to denote a lattice. The familiar $x \cdot y$ will be used to denote the bilinear form, and $x^2 = x \cdot x$ will be called the *norm*. We will only be concerned here with rational-valued forms. Note that calling Λ a finitely generated free Z-module implies that Λ is isomorphic to the module Z^n , where n is the rank, or dimension, of Λ . In other words, algebraically Λ can simply be considered to be Z^n, so each lattice vector x can be represented by a column vector $\vec{x}$ with integer components. Then the bilinear form becomes an $n \times n$ matrix A (called the Gram matrix), where

$$x \cdot y = \vec{x}^T A \vec{y}.$$

There are many different ways to identify Λ with Z^n, and each way involves choosing a different basis $\beta = \{b_1, \ldots, b_n\}$ for Λ . A basis of course is

simply a set of linearly independent (over Z) lattice vectors whose Z-span, $\{m_1 b_1 + \cdots + m_n b_n | m_1, \ldots, m_n \in Z\}$, equals Λ. The Gram matrix A has entries $A_{ij} = b_i \cdot b_j$; these will always be in Q.

A is symmetric, which means it can be expressed as

$$A = B^T G B,$$

where B is an invertible (real) matrix, and where G is a diagonal matrix whose entries are either 0, 1 or -1 (see CAR, pp.5-6). When G (or A) fails to be invertible, Λ is called *singular*. We will only be interested in the case where G is non-singular — i.e. G's diagonal elements must be ± 1 (this implies that our lattices will all be discrete, and their basis vectors will be R-independent, and not merely Z-independent). If G is singular, Λ is called *singular*. Let n_+ (n_-) be the number of +1 (-1) entries. Then $n_+ + n_- = n$, and wlog (="without loss of generality") we can assume $n_+ \geq n_-$. If $n_- > 0$, Λ is said to be an indefinite lattice; if $n_- = 0$, Λ is said to be a positive definite, or Euclidean lattice. Most of the lattices considered here will be Euclidean. When $n_- = 1$, Λ is also called Lorentzian. (The Sylvester law of inertia — CAR — says that n_+ and n_- are well-defined, i.e. independent of the specific choice of B, so these designations are meaningful.)

Thus, there is a second, geometric interpretation of a lattice. Choose any n independent vectors b_k in some space $R^{m,\ell}$ ($R^{m,\ell}$ has the inner product given by the metric tensor $G^{m,\ell} = \mathrm{diag}(\underbrace{+1, +1, \ldots, +1}_{m}, \underbrace{-1, \ldots, -1}_{\ell})$). Then the Z-span of these b_k is an n-dimensional lattice whose bilinear form is induced by the inner product on $R^{m,\ell}$.

Note, however, that these two interpretations are incompatible in the sense that when the lattice is identified with Z^n, the bilinear form is given in general by a non-diagonal matrix A, and when the bilinear form is given by G, the lattice vectors will no longer have integer coefficients. (But see § 3.)

The simplest examples of lattices are the cubic lattices $Z^{m,n}$, consisting of those points in $R^{m,n}$ with integral coordinates (using of course the inner product $G^{m,n}$). These lattices are of dimension $m + n$.

Consider any lattice Λ with basis vectors $b_1, \ldots, b_n$. Suppose their components in some "background space" $R^{m,\ell}$ are $\vec{b_k} = (b_k^1, b_k^2, \ldots, b_k^m | b_k^{m+1}, \ldots, b_k^{m+\ell})$, chosen so that $b_i \cdot b_j = \vec{b_i}^T G^{m,\ell} \vec{b_j}$. Define the generator matrix M by $M_{ij} = b_i^j$,

i.e.

$$M = \begin{pmatrix} b_1^1 & b_1^2 & \cdots & b_1^{m+\ell} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ b_n^1 & b_n^2 & \cdots & b_n^{m+\ell} \end{pmatrix} = \begin{pmatrix} \vec{b_1} \\ \cdot \\ \cdot \\ \cdot \\ \vec{b_n} \end{pmatrix}$$

Then the Gram matrix A, characterizing the induced bilinear form on Λ, is simply $A = MGM^T$. In the special — and by far the most common — case where Λ is Euclidean, we get $A = MM^T$. In both definite and indefinite cases, however, we get:

$$(k_1 b_1 + \cdots + k_n b_n) \cdot (k_1' b_1' + \cdots + k_n' b_n') = k^T A k'.$$

A is an $n \times n$ matrix, while M is $n \times (m + \ell)$.

This hybrid interpretation is the one that will be used thoughout this paper, for the most part. The lattice vectors are uniquely identified by n-tuples of integers, just as in the Z^n interpretation, so the bilinear form is given by a non-diagonal yet symmetric matrix A. But this identification of Λ with Z^n is explicitly recognized by selecting basis vectors, and the bilinear form A is geometrically induced by embedding the basis vectors — and hence Λ — in some $R^{m,\ell}$. Although A doesn't equal some G (except for the cubic lattices), it is derived from both it and the choice of basis vectors. With this understanding having been made, it suffices to characterize a lattice by giving a set of points (albeit one closed under Z-linear combinations) in some $R^{m,\ell}$. Thus the cubic lattice may be simply called $Z^{n,m}$, for its bilinear form could be immediately inferred.

We will almost exclusively be interested in *integral* lattices — i.e. lattices where $x \cdot y \in Z$ for all lattice vectors x and y. In other words, a lattice is integral iff its bilinear (as opposed to quadratic) form is integer-valued. Λ is integral iff A is a Z-matrix (i.e. is a matrix whose entries are all integers). This imposes a strong constraint not only on the norms of the lattice vectors, but on the angles between them.

Unless otherwise stated, assume that Λ denotes an integral and Euclidean lattice. Thus $\Lambda \subset R^m$ and $A = MM^T$. For such lattices, the vectors of norm 1 and 2 are the most interesting, as we shall see. (Of course we always have $n \leq m$.)

An *even* lattice is integral, with the additional property that the norm of every vector in it is even. Λ is even iff A is a Z-matrix whose diagonal entries

are all even. Given any integral lattice Λ, the lattice generated by all the norm 2 vectors in Λ forms an even sublattice of Λ, sometimes called its root lattice. Also, the set of all even-norm vectors in Λ is also an even sublattice of Λ, of the same dimension as Λ (a lattice is like a vector space, except with a field like R replaced by the ring Z . This causes several curious differences — e.g. the possibility of a proper sublattice having the same dimension as the lattice that contains it). When these two even sublattices are equal, we say that Λ is *saturated.* Saturated lattices are readily handled by the machinery of gluing theory, as we shall see later.

An integral lattice which isn't even is called *odd* (so at least one of the diagonal elements of A must be odd). Even lattices seem slightly more important than odd.

Historically, it has often proved convenient to express lattices in the language of quadratic forms. Lattices are studied in number theory largely for this reason. Given any Gram matrix A, construct the quadratic form of n variables $x_1, \ldots, x_n$ by computing the product $x^T A x$. To each quadratic form there is associated a unique lattice in this way, but because each lattice has several equally valid Gram matrices (one for each choice of basis), to each lattice is associated several "integrally equivalent" quadratic forms. Since our picture of lattices will remain geometric, we won't have reason to use the language of quadratic forms.

Lattices also find a spot in group theory when one considers their group of (isometric) symmetries, called their automorphism group. Usually, the automorphism group is meant to include only those symmetries which fix the origin — i.e. translations by lattice vectors are discarded as trivial. It turns out that the automorphism group of integral Euclidean lattices are always finite, but that of indefinite lattices is usually infinite. John Conway used the automorphism group of an important 24-dimensional lattice to find 3 previously undiscovered finite simple groups in 1968 (see §8), and help bring to a close one of mathematics' greatest triumphs. Also, in an unrelated application of lattices, there has been much work concerning translating the study of integral representations of a given finite group G into the study of lattices invariant under that group — i.e. lattices whose automorphism group includes G.

Further applications of lattices can be found in coding theory, computer design, and a number of sphere packing problems (e.g. find the densest packing of identical spheres in R^n — see Figure 3). It is related to the theory

of Lie algebras in a variety of ways (e.g. see §§6-7).

The fundamental region for a lattice, given a basis $b_1, \ldots, b_n$, is the subset of the background space consisting of all points of the form $t_1 b_1 + \cdots + t_n b_n$, for $t_i \in [0, 1)$. It is a building block for the lattice, and when repeated several times fills the whole space (i.e. the whole R-span of the basis vectors) with just 1 lattice point per block (namely, at one of the corners).

A deep hole is a point in the fundamental region whose distance to the closest lattice vector is a global maximum. For example, Z^n has only one deep hole; it has components $(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$. Some lattices though can have more than one deep hole (for example see §8).

Define the *determinant* (also called *discriminant*) of Λ to be $|\Lambda| \equiv |A|$, where A is any Gram matrix of Λ. This is simply the volume-squared of the fundamental region. By Theorem 1 in §3 we see that under a change of basis given by the generator matrices $M \rightarrow M' = UM$, A becomes UAU^T, but $|A|$ becomes $|UAU^T| = |A|$. Thus $|\Lambda|$ is independent of the choice of basis. Shape, but not volume, of the fundamental region is affected by basis transformations.

Define Λ^*, the *dual* of Λ, to be the set of all y in $R\Lambda$, the R-span of the basis vectors, satisfying $y \cdot x \in Z$ for all $x \in \Lambda$. Then Λ^* is an n-dimensional lattice, though in general will only be a rational lattice if Λ is integral. Λ consists of all points in $R\Lambda$ whose contravariant components (relative to b_i) are integers, and Λ^* consists of all points in $R\Lambda$ whose covariant components are integers. Clearly,

$$\Lambda \subseteq \Lambda^* \text{ iff } \Lambda \text{ is integral}$$

in which case Λ^* is denser than Λ. It is possible to prove $(\Lambda^*)^* = \Lambda$. Given a basis b_i for Λ, the dual vectors b_i^* defined by $b_i^* \cdot b_j = \delta_{i,j}$ form a basis for Λ^*, called the dual basis.

2.2 Self-Dual Lattices

Definition 2.2.1 Λ *is called* self-dual, *or* unimodular, *iff* $\Lambda^* = \Lambda$.

We shall exclusively use the term "self-dual" in the following material. The lattices we are most interested in are self-dual. There are both mathematical and physical reasons for this, as we shall see.

Only integral lattices have a hope to be self-dual. In fact, Λ is self-dual iff both it and its dual are integral.

Theorem 2.2.1 *Λ self-dual $\Longrightarrow |\Lambda| = \pm 1$. If $|\Lambda| = \pm 1$ and Λ is integral, then Λ is self-dual.*

> proof: Wlog take Λ to be integral. Choose a basis β of Λ. Let M^* be the generator matrix for Λ^* corresponding to the dual basis of β. Then $\exists$ an $n \times n$ Z-matrix U such that $M = UM^*$, because $\Lambda \subseteq \Lambda^*$, so $I = M^*GM^T = U^{-1}MGM^T = U^{-1}A$. By Corollary 4.2.1, U^{-1} is a Z-matrix iff $|U| = \pm 1$ iff $|\Lambda| = \pm 1$. That is, $\Lambda^* \subseteq \Lambda$ iff $|\Lambda| = \pm 1$. But $\Lambda \subseteq \Lambda^*$ for integral Λ. Thus $\Lambda^* = \Lambda$ iff $|\Lambda| = \pm 1$.

Of course if Λ is Euclidean, we have for integral Λ that Λ is self-dual iff $|\Lambda| = 1$. In other words, self-dual lattices have 1 lattice point per unit volume.

The proof of Theorem 1 can also be used to show that the Gram matrix (relative to the dual basis) for Λ^* is $A^* = A^{-1}$, even for Λ neither integral nor self-dual. Thus for integral Λ we have $\Lambda^* \subseteq \frac{1}{|\Lambda|}\Lambda$.

Odd self-dual lattices will usually be called Type I ; even self-dual lattices will usually be called Type II. The cubic lattices $Z^{m,n}$ are all Type I. There are no trivial examples of Type II lattices (except, if you wish, the 1-point lattice 0). There is no "Type III" lattice, for example, corresponding to those self-dual lattices whose norms are all multiples of 3. In fact, it can be shown that if the norms of the vectors in a self-dual lattice are all multiples of some $k \in Z^+$, then $k = 1$ or 2.

Self-duality is a very strong constraint on lattices, particularly the indefinite ones. $Z^{m,n}$ is the only Type I lattice in the indefinite space $R^{m,n}$ — it is denoted $I_{m,n}$ (this statement only holds of course if both $m, n > 0$). There also is a unique Type II indefinite lattice, called $II_{m,n}$, but it only exists when $m - n$ is a multiple of 8. More precisely, we should say:

Theorem 2.2.2 *Λ is an indefinite Type I lattice of signature (m,n) iff it is* congruent *to $I_{m,n}$, and Λ is an indefinite Type II lattice of signature (m,n) iff $m - n \equiv 0$ (mod 8) and it is* congruent *to $II_{m,n}$ ($II_{m,n}$ will be explicitly given below).*

(See the next section for the definition of congruence). $II_{1,1}$ is given by the Gram matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

For $k \geq 0$ we have $II_{8k+n,n} = II_{1,1}^n \oplus E_8^k$, where by $II_{1,1}^n$ we mean the direct sum

$$\underbrace{II_{1,1} \oplus \cdots \oplus II_{1,1}}_{n}$$

(see §4 for the definition of direct sum), and where E_8 is the unique Type II definite 8-dimensional lattice — it's a root lattice (see §6). These results are all proven in SER, pp.48-58 (Serre and many others use Γ_8 for E_8, and Γ_n for what we'll later call D_n^+).

Example 2.2.1 *Lorentzian Type II lattices must have dimension* $N = 8m + 2$. *They can be defined as the set of all* $x \in R^{N-1,1}$ *satisfying both*
(i) $x \cdot T \in Z$, *and*
(ii) either $x \in Z^{N-1,1}$ *or* $x - T \in Z^{N-1,1}$.
Here, $T = (\frac{1}{2}, \ldots, \frac{1}{2}|\frac{1}{2})$.

The definite case is more complicated, and hence a little richer. All self-dual lattices are known only in dimensions $n \leq 25$, and for reasons I'll give shortly it is doubtful much more progress will be made along these lines. It has been shown that for each dimension only a finite number of self-dual lattices exist, but unfortunately that number quickly reaches unwieldly magnitudes if you go much past $n = 25$.

Again it happens that there is at least 1 Type I lattice in each dimension. And as in the above case, there are Type II lattices only in dimensions which are multiples of 8. (Not surprisingly, the complex Type II lattices, i.e. lattices in C^n rather than R^n, can occur only in dimensions $n \equiv 0 \pmod 4$.) The Type II results are intimately connected: if Λ is a definite Type II lattice, $II_{1,1} \oplus \Lambda$ is a Lorentzian Type II lattice, and so the dimension of Λ must be a multiple of 8. (This argument doesn't imply that there is only one definite Type II lattice in each dimension $n = 8k$; $II_{1,1} \oplus \Lambda = II_{1,1} \oplus \Lambda'$ doesn't give us $\Lambda = \Lambda'$ — but see Theorem 2 of §4). Since D_{8k}^+ is always Type II, we could have instead defined $II_{8k+n,n} = II_{1,1}^n \oplus D_{8k}^+$, even though $D_{8k}^+ = E_8^k$ only when $k = 1$. (Incidently, though different lattices, E_8^2 and D_{16}^+ have the same number of vectors of norm ℓ, for each ℓ — i.e. their theta series are identical.)

Choose any null vector $w \in II_{8k+1,1}$. Then $w^\perp$, the subset of all $u \in R^{8k+1,1}$ satisfying $w \cdot u = 0$, is an $8k+1$-dimensional subspace of $R^{8k+1,1}$ and includes all points in Rw. This means $w^\perp \cap II_{8k+1,1}$ is an $8k+1$-dimensional

sublattice of $II_{8k+1,1}$, containing the points $Zw = \langle w \rangle$. Note that x and $x+\ell w$ have the same norm for $x \in w^\perp \cap II_{8k+1,1}$, so the induced inner product on $\Lambda = (w^\perp \cap II_{8k+1,1})/\langle w \rangle$ is well-defined. Λ is an $8k$-dimensional lattice, both Euclidean and Type II. Depending on the w chosen, this construction will produce every Euclidean Type II lattice.

In 1938, L.J. Mordell proved $E_8 = D_8^+$ was the unique Type II lattice in 8-dimensions. In 1941 E. Witt showed $E_8^2 = E_8 \oplus E_8$ and D_{16}^+ were the only Type II lattices in 16-dimensions, and in 1968 H.-V. Niemeier found all 24 Type II lattices in 24-dimensions (although the most important of these, the Leech lattice, was found in 1965 by John Leech).

In 1957 M. Kneser enumerated all Type I lattices in dimensions $n \leq 16$. Conway and Sloane extended this to $n \leq 23$ in 1982, and in his Ph.D. dissertation in 1984, Borcherds handled $n = 24$ and $n = 25$. See Table 1 for a summary of the results known[1] (Theorem 2.4.1 implies the recursion $a_{n+1} = a_n + b_{n+1} + c_n$; the values of d_n can also be derived from the other columns — see §4 for the definition of indecomposable).

The Minkowski-Siegel "mass" formulae can be used to show these enumerations are complete (apparently the original German is "massformel", which actually means "measure formula", but this mistranslation is now in standard usage). For example:

Theorem 2.2.3 *Let Ω be the set of all (non-congruent) Type II Euclidean lattices of dimension n. Then*

$$\sum_{\Lambda \in \Omega} \frac{1}{|Aut(\Lambda)|} = \frac{|B_k|}{2k} \prod_{j=1}^{k-1} \frac{|B_{2j}|}{4j}$$

where $n = 2k$ is a multiple of 8.

Here, $|\mathrm{Aut}(\Lambda)|$ is the order of the automorphism group of Λ, and where B_k is the kth Bernoulli number. A similar, but more complicated, result holds for Type I Euclidean lattices. There are several standard ways of finding the automorphism groups (we shall see a couple of these in §4 and §7), so Theorem 2 provides a straightforward, if somewhat messy, way of verifying the completeness of the enumerations listed in Table 1.

[1]Table 1 is based on Table 2 2 in CS. The references alluded to here can also be found in CS

Table 1: The n-dimensional Self-dual Euclidean Lattices

Dim. $= n$	Total Number Type I$= a_n$	Number With No Unit Vectors$=b_n$	Total Number Type II$= c_n$	Indecomposable$= d_n$
1	1	0	0	1 (Z^1)
2	1	0	0	0
3	1	0	0	0
4	1	0	0	0
5	1	0	0	0
6	1	0	0	0
7	1	0	0	0
8	1	0	1	0+1 (E_8)
9	2	0	0	0
10	2	0	0	0
11	2	0	0	0
12	3	1	0	1 (D_{12}^+)
13	3	0	0	0
14	4	1	0	1
15	5	1	0	1
16	6	1	2	1+1
17	9	1	0	1
18	13	4	0	4
19	16	3	0	3
20	28	12	0	11
21	40	12	0	12
22	68	28	0	27
23	117	49	0	48
24	273	156	24	154+22
25	665	368	0	367

Note that $|\mathrm{Aut}(\Lambda)| \geq 2$ since $x \to -x$ is always a symmetry. Thus doubling the right-hand side of the formula gives a (crude) lower bound for the number of Type II lattices of dimension n. For example, this gives us $\approx 10^{-9}$ for $n = 8$, $\approx 5 \times 10^{-18}$ for $n = 16$, and $\approx 10^{-14}$ for $n = 24$ (instead of 1, 2 and 24 respectively). But for $n = 32$ it gives (almost certainly) a *very* crude lower bound of 80 million. It seems rather doubtful Niemeier's work will ever be extended.

Similar lower bounds can be found for Type I lattices. For $n = 20$ we get a bound of about 10^{-12} (instead of the actual number of 28). But for $n = 28$ we get about 200, for $n = 29$ we get about 40 000, $n = 30$ about a billion, $n = 31$ a trillion, and $n = 32$ about 10^{17}.

Mathematically, the enumerations of self-dual lattices can be used in a fairly simple manner to find all lattices with other determinants (particularly the smaller determinants). For example, there are 24 lattices of dimension 17 which have determinant 2, and 53 with determinant 3.

One important application of the enumerations for physics is that it enormously constrains the possible gauge groups of various string theories. In fact, the only Yang-Mills groups that can be incorporated in the heterotic string are $\mathrm{Spin}(32)/Z_2$ and $E_8 \times E_8$, corresponding to the only even self-dual 16-dimensional lattices.

2.3 Lattice Equality: Congruence and Similarity

A lattice can be considered as a subset of some R^m with certain properties (i.e. discreteness, closure under Z-linear combinations, and an induced inner product consistent with its Gram matrix). But in some ways this is unsatisfying: for one thing we'd like to think that the background space R^m, though computationally desirable, is hardly intrinsic to the lattice itself. A lattice seems in some fundamental way to be independent of the background space it lies in. I hope to clarify this point in the next couple paragraphs.

Any given lattice, lying in some R^m, can be trivially embedded in each $R^{m'}$, for $m' > m$, by adding $m' - m$ zeroes to the coordinates of each point in the lattice. The reverse, deleting $m - m'$ coordinates, all of which vanish for each point in Λ, will sometimes also be possible. Let $\iota^m_{m'} : R^m \to R^{m'}$ denote this embedding ($m' > m$) or projection ($m' < m$). The Gram matrix is unaffected by $\iota^m_{m'}$. There seems only to be a superficial, easily ignored difference between the lattices Λ and $\iota^m_{m'}(\Lambda)$.

Consider next a global rotation(about the origin) $B : R^m \rightarrow R^m$. Since B leaves inner products unchanged, the Gram matrix is unaffected. Again, the lattices Λ and $B(\Lambda)$ are intimately related — in this case they differ only through a different choice of coordinatization of its basis vectors in R^m, or, alternatively, through an orthogonal change of basis of the background space R^m. Of course, any orthogonal change of basis will work here, but because the metric of R^m must remain Euclidean, non-orthogonal transformations are inadequate (otherwise any lattice could be identified with Z^n).

Both ι and B transform the background space, leaving the Gram matrix and hence the lattice itself untouched. Now, few people would identify a lattice with its basis, and those who do should be quietly dismissed. A lattice is *Z-generated* by the basis β; there is no reason to suppose an alternate basis β' could not be found.

Theorem 2.3.1 *Suppose $\beta = \{b_1, \ldots, b_n\}$ is a basis for a lattice Λ. Then $\beta' = \{b'_1, \ldots, b'_n\}$ is another basis for Λ iff $n' = n$, and the change-of-basis matrix U from β to β', i.e. $b'_i = \sum_{j=1}^{n} U_{ij} b_j$, is a Z-matrix with determinant* ± 1.

This theorem tells us there are several alternate bases, and it characterizes them all. (If Λ was instead a real vector space, U would only have to be an R-matrix with nonzero determinant.) Its proof is simple: the Z-span of β' equals the Z-span of β iff both U and U^{-1} are Z-matrices. Corollary 4.2.1 informs us that this means $|U|$ must be ± 1.

In general, changing the basis of the lattice will change the Gram matrix, but there is little doubt the lattice itself remains unchanged.

This discussion leads us to the concept of congruence.

Definition 2.3.1 *Two lattices Λ and Λ' are called* congruent *if their generator matrices M and M' are related by*

$$M' = \begin{cases} UM\iota^m_{m'}B & \text{if } m \leq m' \\ UMB\iota^m_{m'} & \text{if } m > m' \end{cases}$$

Here U must be a Z-matrix with determinant ± 1, and B is orthogonal. m and m' are the number of columns in M and M', respectively. The matrix $\iota^m_{m'}$ consists of zeroes everywhere, except along the diagonal where it is 1. In addition we must demand that, in the case $m > m'$, the projection operator

$\iota_{m'}^{m}$ only projects out the directions orthogonal to $B(\Lambda)$, or equivalently, that $M = U^{-1}M'\iota_{m}^{m'}B^{-1}$. (If $m = m'$, $\iota_{m'}^{m} = I$.)

It isn't difficult to verify that congruence is an equivalence relation. The Gram matrix $A' = UAU^T$, so $|\Lambda| = |\Lambda'|$. If Λ and Λ' are congruent, then Λ is integral, self-dual, or even iff Λ' is integral, self-dual, or even, respectively. Congruent lattices may, and will, be thought of as the same abstract object coordinatized in different ways. (Lattices, unlike vector spaces, require 2 choices of bases: one for the background, and one for generating the lattice).

See Figures 1 and 4 for several examples of congruences (e.g. Z^2 and $D_2[2]$ are congruent).

Theorem 2.3.2 *Any n-dimensional lattice is congruent to a lattice in* R^n.

This follows because any unit vector can be rotated onto any other unit vector. Of course, this theorem does not suggest that only R^n should be used as a background; in fact the root lattices A_n, E_6 andE_7 (see §6) are most conveniently expressed in spaces with more than the minimum number of dimensions ($m = n + 1$, 8 and 8 are used instead of $n = n$, 6 and 7).

Congruent lattices may, and will for the most part, be considered as equal.

An additional way of constructing one lattice given another is to scale it differently: for each $\lambda \in R$ let $\lambda\Lambda$ denote the lattice $\{\lambda x | x \in \Lambda\}$. Here we have $M \to \lambda M$, $A \to \lambda^2 A$, $|\Lambda| \to \lambda^{2n}|\Lambda|$, and $\Lambda^* \to \frac{1}{\lambda}\Lambda^*$. The lattices Λ and $\lambda\Lambda$ have the same "shape"; the transformation can be (passively) thought of as a mere change in the unit of measurement.

Definition 2.3.2 Λ *and* Λ' *are* similar *lattices, written* $\Lambda \cong \Lambda'$, *if* $\exists\lambda \neq 0$ *such that* $\lambda\Lambda$ *and* Λ' *are congruent.*

The term "equivalent" is also used, but "similar" has a more descriptive geometric flavour. Similarity is an alternate, even weaker interpretation of the equality of lattices. Thus a 'lattice' may be considered to be an equivalence class of either congruent or similar lattices, but we won't bother doing this explicitly.

See Figures 1 and 4 for some examples of similarities (e.g. $Z^2 \cong D_2$).

However, unlike congruence, similarity does not respect self-duality, "integral-ity", or evenness, and so is usually too inclusive to be used to define lattice equality in what follows. Occasionally (e.g. the root lattices) it is

unclear which scale factor λ should be chosen — i.e. which representative of a class of similar lattices should be singled out. The natural convention is to select λ so that $|\Lambda|$ is made as small as possible, and yet the lattice remains integral (of course, this isn't always possible).

The notion of similarity crops up quite naturally in many constructions. The sequence of laminated lattices (see §8) includes many root lattices, all of which are scaled "incorrectly".

Example 2.3.1 *All 1-dimensional lattices are similar to Z.*

Example 2.3.2 *In Figure 1 is shown (among other things) part of the 2-dimensional root lattice A_2, also called (for obvious reasons) the hexagonal lattice. Considering this lattice as being embedded in R^2, a natural basis yields the generator matrix*

$$M = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

and a Gram matrix

$$A = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$$

The hexagonal lattice represented in this way is not integral, since not all of the components of A are integers. The volume/area-squared of the fundamental cell of this lattice is $|A| = \frac{3}{4}$.

More commonly, this lattice is embedded in R^3 with a basis chosen so that

$$M' = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \text{ and } A' = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

This lattice is now integral, with determinant 3.

These lattices are similar, with $\lambda = \sqrt{2}$. The second representation is preferred because in it the generator matrix is simpler, and most importantly, the corresponding lattice is integral (in fact, even).

There are some interesting trivia in which similarity plays a role. Consider, for example, a self-dual lattice Λ (so $\Lambda^* = \Lambda$). Then if $\Lambda' \cong \Lambda$, we must have $\Lambda'^* \cong \Lambda'$, but in general $\Lambda'^* \neq \Lambda'$. This is obvious. Less clear is

converse fails: you can have $\Lambda^* \cong \Lambda$, but have no lattice similar to Λ be self-dual.

Let Λ_{16} denote the 16-dimensional Barnes-Wall lattice, whose generator and Gram matrices are given in Figure 2 (for readability only the nonzero elements in M_{16}, and only half of the symmetric matrix A_{16}, are shown). It is the unique laminated lattice in 16-dimensions (see §8). It's even, and has several special properties: for example, it yields the densest sphere packing known in 16 dimensions. In addition, it satisfies $\Lambda_{16}^* \cong \Lambda_{16}$, but $\Lambda_{16}^* \neq \Lambda_{16}$ — in fact, $|\Lambda_{16}| = 256$. It isn't difficult to show that no lattice similar to Λ_{16} can be self-dual (this would require $(256)^{-\frac{1}{16}} A_{16}$ to be a Z-matrix, where A_{16} is its Gram matrix).

Another counterexample to the converse is the 12-dimensional Coxeter-Todd lattice K_{12}. It is also widely studied, and for instance is the densest packing known in 12 dimensions. $|K_{12}| = 729$, so quite definitely K_{12} isn't self-dual. But again we have that both $K_{12}^* \cong K_{12}$, and no lattice similar to K_{12} is self-dual.

Far simpler examples include A_2 and D_4. (Any 1-dimensional lattice is similar to its dual, but is also similar to the self-dual lattice Z, by Example 1.) To see that $A_2^* \cong A_2$, note that (using the notation of Example 2)

$$A^* = A^{-1} = \frac{4}{3} \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} = \frac{2}{3} A'$$

and use Theorem 3.

On the other hand, if an (integral) lattice is congruent to its dual, then it equals its dual.

We shall conclude this section with two simple examples of congruence and similarity.

Theorem 2.3.3 *Let Λ and Λ' be two lattices with Gram matrices A and A'. Then:*

1. *if $A = A'$, Λ and Λ' are congruent;*

2. *if $A = \lambda^2 A'$, Λ and Λ' are similar with scale factor λ.*

The proof of 1. is geometrically obvious; 2. follows immediately from 1.

Figure 2: The Barnes-Wall Lattice

(only the nonzero elements in M_{16}, and only half of the symmetric matrix A_{16}, are shown)

$$M_{16} = \frac{1}{\sqrt{2}} \begin{pmatrix}
4 & & & & & & & & & & & & & & & \\
2 & 2 & & & & & & & & & & & & & & \\
2 & & 2 & & & & & & & & & & & & & \\
2 & & & 2 & & & & & & & & & & & & \\
2 & & & & 2 & & & & & & & & & & & \\
2 & & & & & 2 & & & & & & & & & & \\
2 & & & & & & 2 & & & & & & & & & \\
2 & & & & & & & 2 & & & & & & & & \\
2 & & & & & & & & 2 & & & & & & & \\
2 & & & & & & & & & 2 & & & & & & \\
2 & & & & & & & & & & 2 & & & & & \\
1 & 1 & 1 & 1 & & 1 & & 1 & 1 & & & 1 & & & & \\
 & 1 & 1 & 1 & 1 & & 1 & & 1 & 1 & & & 1 & & & \\
 & & 1 & 1 & 1 & 1 & & 1 & & 1 & 1 & & & 1 & & \\
 & & & 1 & 1 & 1 & 1 & & 1 & & 1 & 1 & & & 1 & \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}$$

$$A_{16} = \begin{pmatrix}
8 & & & & & & & & & & & & & & & \\
4 & 4 & & & & & & & & & & & & & & \\
4 & 2 & 4 & & & & & & & & & & & & & \\
4 & 2 & 2 & 4 & & & & & & & & & & & & \\
4 & 2 & 2 & 2 & 4 & & & & & & & & & & & \\
4 & 2 & 2 & 2 & 2 & 4 & & & & & & & & & & \\
4 & 2 & 2 & 2 & 2 & 2 & 4 & & & & & & & & & \\
4 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & & & & & & & & \\
4 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & & & & & & & \\
4 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & & & & & & \\
4 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & & & & & \\
2 & 2 & 2 & 2 & 1 & 2 & 1 & 2 & 2 & 1 & 1 & 4 & & & & \\
0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 2 & 4 & & & \\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 2 & 4 & & \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 2 & 4 & \\
2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2
\end{pmatrix}$$

2.4 Direct Sums

There are 3 basic ways of building higher dimensional lattices up from smaller ones (a fourth one, direct product, is possible but never seems to be used):

(i) Direct sums (introduced in this section);

(ii) Gluings (introduced in §7); and

(iii) Lamination (introduced in §8).

Let $\Lambda_1, \ldots, \Lambda_k$ be lattices of dimension $n_1, \ldots, n_k$. Consider the set $\Lambda = \{(x_1, \ldots, x_k) \mid x_i \in \Lambda_i\}$. For any 2 points $x = (x_1, \ldots, x_k)$ and $y = (y_1, \ldots, y_k)$ in Λ, define $x+y = (x_1+y_1, \ldots, x_k+y_k)$, $x \cdot y = x_1 \cdot y_1 + \cdots + x_k \cdot y_k$, and for $\lambda \in R$, $\lambda x = (\lambda x_1, \ldots, \lambda x_k)$.

Obviously this makes Λ an $(n_1 + \cdots + n_k)$-dimensional lattice. It is called the *direct sum* of the *components* Λ_i, and is denoted by $\Lambda_1 \oplus \cdots \oplus \Lambda_k$. Note that it is an orthogonal sum — i.e., loosely speaking $\Lambda_i \perp \Lambda_j$ for $i \neq j$. Also, any $x \in \Lambda$ can be uniquely written as $x = \sum_{i=1}^{k} x_i$, where $x_i \in \Lambda_i$ (again, the notation adopted here is deliberately naive, chosen for simplicity).

A basis for Λ consists of the union of bases of each component. For this basis,

$$M = \begin{pmatrix} M_1 & & 0 \\ & M_2 & \\ & & \ddots \\ 0 & & M_k \end{pmatrix} \text{ and } A = \begin{pmatrix} A_1 & & 0 \\ & A_2 & \\ & & \ddots \\ 0 & & A_k \end{pmatrix}$$

using obvious notation. Thus Λ is integral, even or self-dual iff each component is. The minimal norm μ of Λ equals the smallest of the minimal norms μ_i of the components, and $|\Lambda| = \prod |\Lambda_i|$. We will write Λ^ℓ for $\underbrace{\Lambda \oplus \cdots \oplus \Lambda}_{\ell}$.

Definition 2.4.1 *Call* Λ indecomposable *if* $\Lambda = \Lambda_1 \oplus \Lambda_2 \Longrightarrow \Lambda_1$ *or* Λ_2 *is the zero lattice, i.e. if* Λ *cannot be written as the direct sum of proper sublattices.*

Indecomposable lattices are the basic building blocks (see column 5 of Table 1). Direct sums can be defined for vector spaces; the only indecomposable vector space is (up to isomorphism) R. Lattices are much more interesting in this respect. For example, apart from the Z^k, all root lattices (see §6) are indecomposable. D_4^+ (see §7), however, is not (in fact, $D_4^+ = Z^4$). The only indecomposable integral lattice containing unit vectors is Z. Indeed, it

is possible to use direct sums to characterize all integral lattices with unit vectors.

Theorem 2.4.1 *Any integral lattice Λ can be uniquely expressed as the direct sum $\Lambda = Z^k \oplus \Lambda'$, where Λ' contains no unit vectors, and where Λ has exactly $2k$ unit vectors. Λ is self-dual iff Λ' is.*

Λ' is called the reduced version of Λ. This result is exploited in various enumerations — for example, to find all self-dual lattices it is sufficient to find those containing no unit vectors (see column 5 of Table 1). For smaller dimensions (less than 20 or so), well under half of all Type I lattices are of this form. A slightly weaker theorem (Witt's Theorem in §6) applies to vectors of norm 2 (with the role of direct sums being taken by gluings). Incidently, all lattice equalities expressed in this section are, properly speaking, actually congruences.

proof of Theorem 1: The proof is analogous to the Gram-Schmidt orthogonalization process.

Let $\{b_1, \ldots, b_n\}$ be a basis for Λ. Let $2k$ be the number of unit vectors in Λ (this number is even because $u \cdot u = (-u) \cdot (-u)$) and let $u_1, \ldots, u_k$ be k linearly independent unit vectors in Λ (it is sufficient to ensure $u_i \neq \pm u_j$ for all $i \neq j$). Then for $i \neq j$, $u_i \cdot u_j$ must be an integer (as Λ is integral), and also must satisfy $-1 = -u_i^2 \cdot u_j^2 < (u_i \cdot u_j)^2 < u_i^2 \cdot u_j^2 = 1$ because of linear independence. Therefore, $u_i \cdot u_j = 0$ whenever $i \neq j$.

Let $b_i' = b_i - \sum_{j=1}^{k} (b_i \cdot u_j) u_j$. Then $b_i \cdot u_j = 0 \ \forall i, j$. Note that if $\sum \alpha_i u_i + x = \sum \beta_j u_j + y$, where $\alpha_i, \beta_j \in Z$ and $x, y \in \langle b_1', \ldots, b_n' \rangle$, then dotting this with u_ℓ gives $\alpha_\ell = \beta_\ell$ for each j, implying as well that $x = y$. Also, the Z-span $\langle u_1, \ldots, u_k, b_1', \ldots, b_n' \rangle$ must equal Λ. Thus, $\Lambda = Z \oplus \cdots \oplus Z \oplus \langle b_1', \ldots, b_n' \rangle = Z^k \oplus \Lambda'$, where there is one Z for each u_i, and where $\Lambda' = \langle b_1', \ldots, b_n' \rangle$.

Of course, Λ' can contain no unit vectors, for such a vector (or its negative) would have had to be contained in our original list $u_1, \ldots, u_k$ of unit vectors.

In GO we find the statement: "A euclidean lattice has a unique decomposition into a direct sum of indecomposable sublattices ...". They don't

prove this, and unfortunately don't clear up a possible ambiguity concerning exactly what this means. There are two possibilities:

(a) (the weaker sense) $\Lambda_1 \oplus \cdots \oplus \Lambda_k = \Lambda = \Lambda'_1 \oplus \cdots \oplus \Lambda'_{k'}$, where Λ_i, Λ'_j are all indecomposable lattices, implies $k = k'$ and $\Lambda_i = \Lambda'_{\sigma_i}$ for some permutation $\sigma \in S_k$ (again, equality here denotes congruence).

For example, this is trivially satisfied by the vector space $R^n = R \oplus \cdots \oplus R$. These n components R correspond to n 1-dimensional orthogonal subspaces of R^n. Similarly, each of the Λ_i correspond to a sublattice of Λ congruent to Λ_i, in such a way that all of these sublattices are mutually orthogonal.

However, note that the decomposition referred to in (a) can be far from unique in another sense. For example, the 2 orthogonal subspaces in R^2 can be chosen to be $\{(x,0) \mid x \in R\}$ and $\{(0,y) \mid y \in R\}$, or $\{(x,x) \mid x \in R\}$ and $\{(y,-y) \mid y \in R\}$.

(b) (the stronger sense) The same as above, except both Λ_i and Λ'_{σ_i} must correspond to the same sublattice in Λ. In other words, consider only the decompositions $\Lambda_1 \oplus \cdots \oplus \Lambda_k$ in (a) whose Λ_i are orthogonal sublattices of Λ (and not merely congruent to those sublattices). Then for uniqueness in this stronger sense to hold, there must be a permutation σ such that $\Lambda_i = \Lambda'_{\sigma_i}$, where now equality is not merely congruence.

For example, R^n doesn't decompose uniquely in this sense. GO probably meant this stronger interpretation. In any case, for lattices (unlike vector spaces), the uniqueness of (b) holds.

Theorem 2.4.2 *A (Euclidean) integral lattice has a unique decomposition into a direct sum of indecomposable sublattices, in the strong sense of (b).*

> proof: First we shall construct one such decomposition, and then we shall show it is unique.
>
> Let $L_k = \{x \in \Lambda \mid x^2 = k\}$. Note that each L_k is a finite set (i.e. $|L_k| < \infty$). Choose any basis $\{b_1, \ldots, b_n\}$ for Λ, and find an N satisfying $N \geq b_i^2$ for all i.
>
> What we will do is run through the finitely many vectors in $L = \cup_{k=1}^{N} L_k$, discarding some of these, and partitioning the remainder into a number of sets C_i. The C_i generate the components Λ_i of the decomposition. The discarded vectors are hybrids, sums of vectors from more than one component.

Enumerate the vectors in L in this way: let $L_1 = \{v_\ell \mid \ell = 1, \ldots, |L_1|\}$, $L_2 = \{v_\ell \mid \ell = |L_1| + 1, \ldots, |L_1| + |L_2|\}$, ..., and $L_N = \{v_\ell \mid \ell = \sum_{k=1}^{N-1} |L_k| + 1, \ldots, \sum_{k=1}^{N} |L_k|\}$.

For convenience write $v_\ell \cdot C_i = 0$ if $v_\ell \cdot v = 0 \; \forall v \in C_i$. Otherwise write $v_\ell \cdot C_i \neq 0$. Start the following construction with each C_i empty.

Place v_1 in C_1, and set $m = 1$ (m is the number of C_i currently nonempty). Proceed recursively for $\ell = 2, 3, \ldots, \sum_{k=1}^{N} |L_k|$.

There are 3 possibilities:

1. if $v_\ell \cdot C_i = 0$ for all $i = 1, \ldots, m$, then let $C_{m+1} = \{v_\ell\}$ and increment m;
2. if $v_\ell \cdot C_i = 0$ for all $i = 1, \ldots, m$, $i \neq j$, and if $v_\ell \cdot C_j \neq 0$, then put v_ℓ in C_j; and
3. otherwise do nothing (i.e. discard v_ℓ).

This recursion will produce m sets C_i. It isn't difficult to show $\Lambda = \langle C_1 \rangle \oplus \cdots \oplus \langle C_m \rangle$ is one of the desired decompositions, but it is unnecessary to directly establish this. Since the existence of such decompositions is not in doubt, it is sufficient to prove that any of these decompositions must equal the one constructed above.

Suppose $u \in \Lambda$ does not lie in a component of a given decomposition of Λ. Then $\exists v, w \in \Lambda$ such that $v \cdot w = 0$ and $0 < u \cdot v = v^2 < u^2$, and $0 < u \cdot w = w^2 < u^2$. This simple result is the central fact of this proof. For example, it immediately follows from this that v_1 must be in some component.

Consider a decomposition $\Lambda = \Lambda'_1 \oplus \cdots \oplus \Lambda'_{m'}$ into indecomposable sublattices Λ'_i of Λ. Let $C'_i = L \cap \Lambda'_i$. The above recursion, followed step by step, shows $m' = m$ and (up to a permutation of the indices) $C_i = C'_i$. This then implies (up to this same permutation) $\Lambda_i = \Lambda'_i$, because a basis of Λ (namely $\{b_1, \ldots, b_n\}$) is contained in L.

Thus the recursively given decomposition is the unique one.

Although this theorem was stated and proved for integral lattices, a slight rewording of the proof establishes it for any (Euclidean) lattice. However, the proof relied much more heavily on the assumption that Λ was Euclidean — indeed, uniqueness (even in the weaker sense) fails for indefinite (e.g. Lorentzian) lattices.

This can be shown in many ways. For example, the uniqueness of $II_{8k+1,1}$ implies that $II_{8k+1,1} = II_{1,1} \oplus \Lambda$ for any $8k$-dimensional (Euclidean) Type II lattice (there is more than 1 $8k$-dimensional (Euclidean) Type II lattice, $\forall k > 1$).

The main consequence of Theorem 2 concerns automorphisms.

Theorem 2.4.3 *Let $\Lambda_1^{n_1} \oplus \Lambda_2^{n_2} \oplus \cdots \oplus \Lambda_k^{n_k}$ be a decomposition of a Euclidean integral lattice Λ into indecomposable lattices Λ_i, where for $i \neq j$, Λ_i and Λ_j are not congruent. Then the automorphism group $Aut(\Lambda) = (Aut\Lambda_1)^{n_1} \times S_{n_1} \times \cdots \times (Aut\Lambda_k)^{n_k} \times S_{n_k}$.*

For example, $\mathrm{Aut}(\Lambda)$ has $|\mathrm{Aut}(\Lambda')|2^k k!$ elements, where $\Lambda' \oplus Z^k$ is the decomposition of Theorem 1. This theorem tells us that it suffices to know the automorphisms of indecomposable lattices only.

Theorem 3 also fails in general for indefinite lattices. For example, $\mathrm{Aut}(II_{1,1})$ is finite, so if it applied to $II_{25,1}$, $|\mathrm{Aut}(\Lambda)|$ would be equal for all Niemeier lattices Λ. But this is false: $|\mathrm{Aut}(\Lambda)|$ varies from around 2×10^7 to 4×10^{23}. Also, the automorphism groups of $II_{25,1}$ and most other indefinite lattices are infinite, unlike those of $II_{1,1}$ and all Euclidean lattices.

2.5 Every Lattice is a Sublattice of Some Z^m

(In this section we will study Euclidean lattices, but similar results apply to indefinite ones). A lattice consists of a set of points "isomorphic" to Z^n, and a positive definite binary form. It can be interpreted to consist precisely of the points in Z^n, but with a non-Euclidean norm given by this form. Another, more geometric interpretation seems to be as a set of points spanning an n-dimensional subspace of R^m, whose norm is induced by the Euclidean norm of R^m (see §1). Here, $m \geq n$ — for example, A_n is usually defined to be contained in R^{n+1} (see Table 2). Though it is easy to prove that any n-dimensional integral lattice is congruent (and not merely similar) to a lattice in R^n (this is Theorem 2 in §3), the basis and hence the generator matrix may take a simpler form in a background space of dimension $m > n$.

An interesting problem, however one unrelated to what follows, concerns whether all lattices can be embedded in Z^m in some sense. (Of course, Z^m denotes the lattice with Euclidean norm, whose basis is the standard basis $\{e_i\}$ of R^m.) Alternatively, this amounts to asking if each Gram matrix A can be factored into *integral* generator matrices M (M is $n \times m$, so may not be square).

Conjecture 1 *Any n-dimensional integral lattice is a sublattice of Z^n, where "is" here is shorthand for "is congruent to".*

Note that here $m = n$. This conjecture would imply that to each Gram matrix there is a square integral generator matrix M such that $A = MM^T$.

> Counterexample 1: This conjecture is very false: the one-dimensional lattice $A_1 = \sqrt{2}Z$ is one of several counterexamples. (See the following section for the definition of the root lattices A_n).

There are two natural directions to proceed from here.

Conjecture 2 *Any n-dimensional integral lattice is similar (e.g. after appropriate scaling) to a sublattice of Z^n.*

For example, here $\sqrt{2}Z$ would simply be rescaled to Z. Again note $m = n$. In this case M is also integral and square, but it is sufficient that A merely equal $\lambda^2 MM^T$ for some λ. Conjecture 2 also fails. As we shall see below, A_2 and hence A_n (for $n \geq 2$) simply cannot be embedded in Z^2 and Z^n, respectively, regardless of the scaling factor. Z^{n+1} is required, in which case the simple roots $e_i - e_{i+1}$ form a basis.

> Counterexample 2: We know A_2 has a basis consisting of 2 independent vectors of norm 2 (usually these are taken to be $e_1 - e_2$ and $e_2 - e_3$). Let this basis be mapped (under the similarity transformation) to the vectors $ae_1 + be_2$ and $ce_1 + de_2$ in Z^2. Then $a, b, c, d \in Z$. Wlog we can assume a, b, c, d don't have a common divisor (otherwise absorb it in the scale factor λ). Then we have
>
> $$a^2 + b^2 = c^2 + d^2 = 2\lambda^1 \text{ (call this equation (a))}$$
>
> $$\text{and } ac + bd = \lambda^2 \text{ (call this equation (b)).}$$

Now (b) implies $\lambda^2 \in Z$, so (a) gives us $a \equiv b$, $c \equiv d \pmod 2$. But $\text{odd}^2 + \text{odd}^2 \equiv 2 \pmod 4$ and $\text{even}^2 + \text{even}^2 \equiv 0 \pmod 4$, so we must have $a \equiv b \equiv c \equiv d \equiv 1 \pmod 2$ (by hypothesis, all coefficients cannot be even). But sticking this into (a) implies λ^2 is odd, and sticking this into (b) implies λ^2 is even. This contradiction implies no such similarity can be found, and Conjecture 2 cannot be true.

Conjecture 3 *Any n-dimensional integral lattice is congruent to a sublattice of some Z^m, for $m \geq n$.*

So here we're fixing λ at 1, but allowing m to vary (so M need no longer be square). For example, A_n can be embedded in this way in Z^{n+1}, and $\sqrt{k}Z$ can be embedded in Z^k.

This conjecture also falls short.

Counterexample 3: Consider the E_6 lattice (E_7 or E_8 will work as well) with a basis $\{r_i\}$ of simple roots. (See the following section, and in particular Table 2, for the definitions of these lattices and a list of their simple roots). All these basis vectors have norm 2. In any Z^m, the only vectors with norm 2 are of the form $\pm e_i \pm e_j$, for $i \neq j$.

Wlog let $e_1 - e_2$ correspond to r_2. $r_2 \cdot r_3 = -1$, so r_3 "contains" either e_1 or e_2 but not both. Wlog say $r_3 = e_2 - e_3$.

$r_4 \cdot r_3 = -1$ so either r_4 contains e_2, in which case $r_4 = -e_1 - e_2$ (since $r_4 \cdot r_2 = 0$), or r_4 contains e_3, in which case we can take $r_4 = e_3 - e_4$. In the first case $r_5 \cdot r_4 = -1$ implies r_5 contains exactly one of e_1 or e_2, contradicting $r_5 \cdot r_2 = 0$. Thus we must have $r_4 = e_3 - e_4$.

Now look at r_6. From $r_6 \cdot r_3 = -1$ and $r_6 \cdot r_2 = r_6 \cdot r_4 = 0$ we get either $r_6 = -e_1 - e_2$ (in which case r_1 cannot be found) or $r_6 = e_3 + e_4$ (in which case r_5 can't be found).

Therefore E_6 provides us with a counterexample to Conjecture 3.

The next, and final, hope for a general theorem asserting the possibility of embedding each integral lattice in Z^m turns out to be true.

Theorem 2.5.1 *Any n-dimensional integral lattice is similar to a sublattice of some Z^m, for m sufficiently big, and where the scaling factor λ may be chosen to be the reciprocal of a positive integer.*

In other words, we can symbolically write $\Lambda \subseteq \lambda Z^m$, for any integral lattice Λ. An alternate formulation is given in Theorem 2 at the end of this section.

Due to the counterexamples given above, this theorem seems to be the strongest result we can hope for. Roughly speaking, the extra dimensions are needed to remove square roots from the numerators of the elements of a generator matrix. The scale factor is needed to remove their common denominator.

Proof of Theorem 1: Let A be the Gram matrix for the lattice, using the basis $\{r_k\}$. Proceeding in a manner analogous to the Gram-Schmidt process, we shall find orthogonal vectors s_k, $k = 1, \ldots, n$, each with integral norm, satisfying

$$\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = B \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix},$$

where B is a lower triangular matrix with rational number entries. This can be done as follows:

Recursively define $A^{(1)} = A$, $s_k^{(1)} = r_k$ and for $\ell = 1, \ldots, n$ let

$$r_{(k)}^{(\ell)} = \begin{cases} 0 & \text{for } k < \ell \\ A_{\ell k}^{(\ell)} s_\ell^{(\ell)} & \text{otherwise} \end{cases}, \quad s_k^{(\ell+1)} = \begin{cases} 0 & \text{for } k < \ell \\ A_{\ell\ell}^{(\ell)} s_k^{(\ell)} - r_{(k)}^{(\ell)} & \text{otherwise} \end{cases},$$

and finally let $A^{(\ell+1)}$ be the Gram matrix for $\{s_1^{(\ell+1)}, \ldots, \{s_n^{(\ell+1)}\}$. Each of these Gram matrices is integral. If we set $\lambda_\ell = A_{\ell\ell}^{(\ell)}$, then

$$r_k = \frac{r_{(k)}^{(1)}}{\lambda_1} + \frac{r_{(k)}^{(2)}}{\lambda_1 \lambda_2} + \cdots + \frac{r_{(k)}^{(n)}}{\lambda_1 \cdots \lambda_n}$$

is an othogonal decomposition of the basis vectors. Setting $s_k = s_k^{(k)}$ and

$$B_{k\ell} = \begin{cases} 0 & \text{for } k < \ell \\ \frac{A_{\ell k}^{(\ell)}}{\lambda_1 \cdots \lambda_\ell} & \text{otherwise} \end{cases},$$

we get the desired expression for the r_k in terms of that for s_ℓ.

The fact that B is only a rational matrix can be corrected by factoring out a common denominator from all of its entries — i.e. by choosing a scale factor $\lambda = \frac{1}{\lambda_1 \cdots \lambda_n}$. It now suffices to find an expression for the s_k in terms of the standard basis of some Z^m. It doesn't matter how this is done, provided their components are only integers.

Note first that s_k has length (as opposed to norm) $\sqrt{\lambda_k}$, so to guarantee that our generator matrix has only integer entries we should write s_1 as the sum $e_1 + \cdots + e_{\lambda_1}$, s_2 as $e_{\lambda_1+1} + \cdots + e_{\lambda_1+\lambda_2}$, etc. Thus we have $m = \lambda_1 + \cdots + \lambda_n$. (Of course, we cannot in general simply absorb these radicals in the scale factor because these radicals need not be equal.)

Theorem 2.5.2 *Let A be any positive definite matrix (so it must be square and symmetric). Then there exists a Z-matrix M (not necessarily square) and a positive integer λ such that*

$$\lambda^2 A = MM^T.$$

(This is an immediate corollary to Theorem 1. This is also the best result possible — in general you need a scale factor and M to be non-square in order to guarantee M only has integer components.)

An almost identical proof establishes Theorem 1 for identical lattices, implying

Theorem 2.5.3 *Let A be any symmetric matrix (so it must be square). Then there exists a Z-matrix M (not necessarily square), a positive integer λ, and a matrix G of the form diag(+1,...,+1,-1,...,-1) such that*

$$\lambda^2 A = MGM^T.$$

2.6 Root Lattices

The theory of Lie groups and algebras is certainly among the most elegant and consequential of all mathematical theories. Its influence is felt in areas such as gauge theory in physics (e.g. superstrings) and the classification of the finite simple groups (16 of the 18 infinite families of finite simple groups are of Lie type). And in this and the following section its significant applications to lattice theory will be presented.

To every complex semi-simple Lie algebra there is associated a root system, i.e. a set of vectors $\{\alpha_i\}$ (called roots) satisfying various properties (e.g. $\frac{2\alpha_i \cdot \alpha_j}{\alpha_i \cdot \alpha_i} \in Z$). A basis for it can be found — these basis vectors are called simple roots. A convenient way of graphically representing a set of simple roots is with a Dynkin diagram: to each simple root there is associated a node in the diagram, and 2 nodes are connected by 0, 1, 2 or 3 (sometimes directed) segments depending on the dot product of the corresponding simple roots. All possible connected Dynkin diagrams are known — they correspond to the complex simple Lie algebras A_n, B_n, C_n, D_n, E_6, E_7, E_8, F_4, and G_2 (all other Dynkin diagrams are simply unions of these).

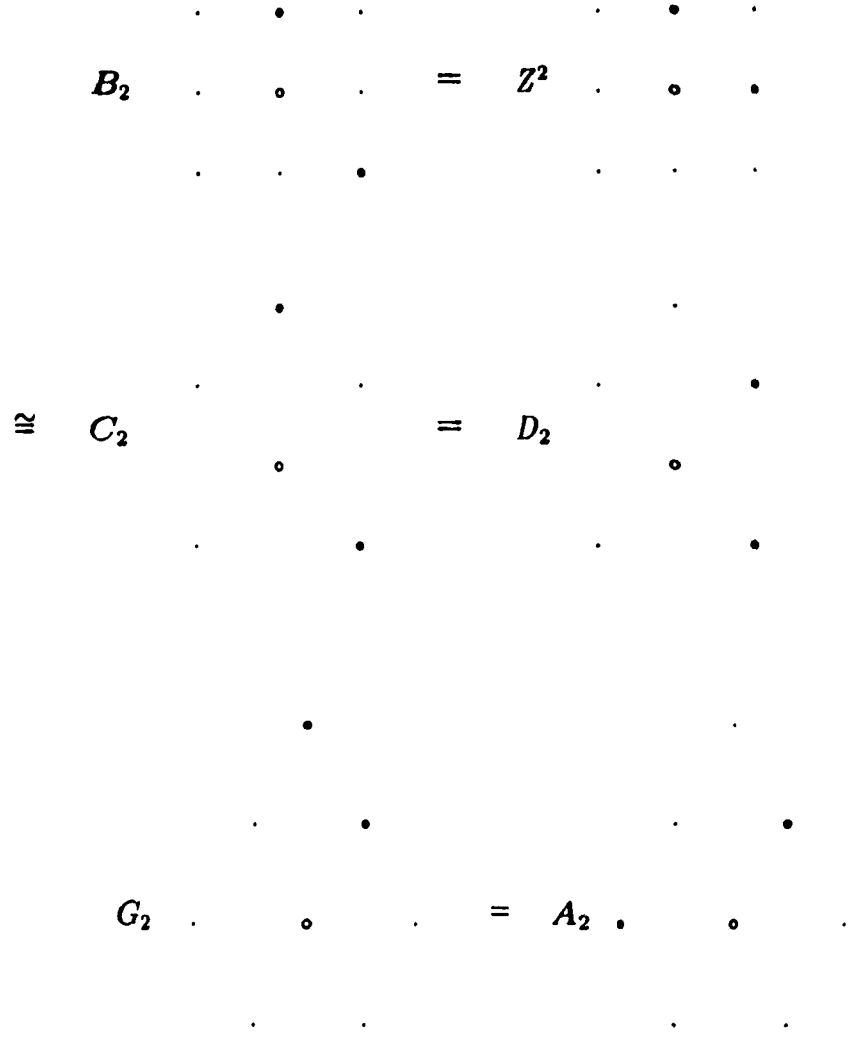
By a root lattice of some Lie algebra we simply mean the lattice generated by the simple roots of the Lie algebra. The simple roots are known up to a global rotation and global scale factor only, but we shall fix them by adopting the conventions of Bourbaki (see BOU, pp. 250-262). The dimension of the root lattice, i.e. the number of simple roots, is the rank of the Lie algebra and is the value of the subscript (e.g. the root lattice for A_n has dimension n). We will use the same symbol to denote the root lattice and the Lie algebra (no complications should result).

See Figure 1 for the 2-dimensional root lattices (the origin is labelled with an "o", and the simple roots with a bullet). As can be seen there, many of the root lattices are congruent ($=$) or similar ($\cong$) to one another.

Theorem 2.6.1 *$B_n = Z^n$, $G_2 = A_2$, $C_n = D_n$, and $F_4 \cong D_4$. That is, it is sufficient to consider only the simply-laced Lie algebras (i.e. the ones whose simple roots all have equal length). In addition, $A_1 = D_1 \cong Z$, $D_2 = D_1^2 \cong Z^2$, and $D_3 = A_3$. These completely exhaust the similarities/congruences between root lattices.*

proof: The similarities/congruences involving A_1, D_1, D_2 and D_3 are all obvious when one considers the corresponding Dynkin

Figure 1: The Two-Dimensional Root Lattices



diagrams.

case (1): $B_n = Z^n$

Let e_k, $k = 1, \ldots, n$ be the standard basis for Z^n. In terms of this basis, the simple roots α_k of B_n are $\alpha_1 = e_1 - e_2, \ldots, \alpha_{n-1} = e_{n-1} - e_n, \alpha_n = e_n$ (according to Bourbaki). Thus B_n has generator matrix

$$M' = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \end{pmatrix},$$

where we have displayed only the non-zero entries. Z^n of course has $M = I$. Setting $\lambda = 1$, $B = I$ and $U = M'$, and noting that $|U| = 1$ (since M' is upper triangular) we see that $M' = \lambda U M B$, so the root lattice B_n is congruent to Z^n. (In fact, here we have that the span of the α_k equals the span of the e_k, so the transition from the α_k to the e_k amounts merely to a change of basis. Thus B_n and Z^n are equal here to an extent even greater than congruence, but this is more due to a fortuitous choice of basis vectors than to some profound relationship between B_n and Z^n.) (We could ignore the projections/embeddings $\iota_m^{m'}$ because we have here $m = m' (= n)$).

case (2): $G_2 = A_2$

Bourbaki gives the simple roots of G_2 to be $\alpha_1' = e_1 - e_2$, $\alpha_2' = -2e_1 + e_2 + e_3$. He gives the simple roots of A_2 to be $\alpha_1 = e_1 - e_2$, $\alpha_2 = e_2 - e_3$. (Here $m = m' = 3$).

$$M' = \begin{pmatrix} 1 & -1 & 0 \\ -2 & 1 & 1 \end{pmatrix}, \; M = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix},$$

$$\Longrightarrow \lambda = 1, \; B = I, \; U = \begin{pmatrix} 1 & 0 \\ -2 & -1 \end{pmatrix}.$$

Since $|U| = -1$, G_2 and A_2 are congruent.

The root system for G_2 consists of all vectors in the A_2 lattice with norms 2 and 6 — the smallest 2 norms in A_2.

case (3): $C_n = D_n$

The simple roots for C_n can be chosen to be $\alpha'_k = e_k - e_{k+1}$ for $k = 1, \ldots, n-1$, and $\alpha'_n = 2e_n$. Those for D_n are $\alpha_k = e_k - e_{k+1}$ for $k = 1, \ldots, n-1$, and $\alpha_n = e_{n-1} + e_n$.

$$M' = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 2 \end{pmatrix}, M = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & 1 & 1 \end{pmatrix},$$

$$U = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & 1 & -1 \end{pmatrix},$$

$\lambda = 1$, and $B = I$, so C_n is congruent to D_n.

case (4): $F_4 \cong D_4$

The simple roots for F_4 are $\alpha'_1 = e_2 - e_3$, $\alpha'_2 = e_3 - e_4$, $\alpha'_3 = e_4$, $\alpha'_4 = \frac{1}{2}(e_1 - e_2 - e_3 - e_4)$. Thus

$$M' = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}, M = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

$$U = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

and $\lambda = \frac{1}{\sqrt{2}}$. Then B is orthogonal, $|U| = -1$, and $M' = \lambda UMB$, so these lattices are similar (but not congruent). The presence of a scale factor, and the realization that the F_4 root lattice represented in this way fails to be integral, suggests that this choice of

simple roots for F_4 isn't the most natural: $\alpha_k'' = \sqrt{2}\alpha_k'$ may look messy, but (at least in the context of lattice theory) is a better choice.

That these exhaust all possible similarities/congruences follow easily by computing determinants (see Table 2), for example.

If we hadn't adopted Bourbaki's selections, Theorem 1 would still be valid, provided we replaced congruences everywhere with similarities. In this sense, similarity is coordinate-independent, unlike congruence, and especially unlike pointwise equality ($B = I, \lambda = 1$), which occurred here in cases (1), (2) and (3).

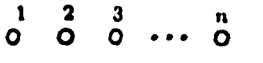
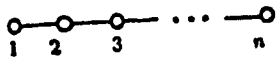
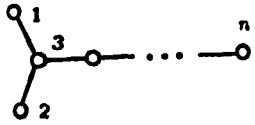
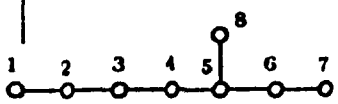
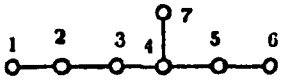
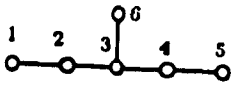
Theorem 1 allows us to completely ignore the multiply-laced root lattices in the following pages. The root lattices A_n, D_n, E_6, E_7, and E_8 all have minimum norm $\mu = 2$ (minimum norm is the smallest non-zero norm), all are integral, all are even and all are indecomposable. (Z^n is odd, has $\mu = 1$, and is only indecomposable if $n = 1$). Only Z^n and E_8 are self-dual, however. The relevant features of the root lattices are given in Table 2. Bourbaki's choice of norm 2 for the simple roots of these simply-laced algebras is the smallest possible choice for which the root lattices are integral (e.g. $\frac{1}{\sqrt{2}}A_n$ isn't integral). (In Table 2 slightly different simple roots were chosen so as to agree with the notation in CS.) Provided we adopt the range restrictions of n given in Table 2, all of these root lattices are distinct, in either sense of §3.

From the Dynkin diagram (also given in Table 2) we can read off the Gram matrix. Each node in the diagram represents a simple root (of norm 2). Two nodes are connected iff the dot product of the corresponding simple roots (= basis vectors) is -1; otherwise the simple roots are orthogonal. Thus the Gram matrix has 2's down the diagonal, and only -1's and 0's scattered off the diagonal.

The Cartan matrix of a root system is the $n \times n$ Z-matrix whose entries are $\frac{2\alpha_i \cdot \alpha_j}{\alpha_i \cdot \alpha_i}$, for the simple roots $\alpha_i, i = 1, \ldots, n$. For simply-laced systems, normalized as we have them normalized, this is precisely the Gram matrix of the corresponding root lattice, relative to the basis of simple roots. Thus we have Gram matrix = Cartan matrix here. Similarly, the weight lattice is just the dual lattice.

Let Λ be any integral lattice, and Λ_R be the sublattice generated by its vectors of norm ≤ 2 (these vectors are often called root vectors of Λ, for

Table 2: The Root Lattices

Symbol	Congru.	Dynkin Diagram	Basis Vectors	Det
Z^n, $n \geq 1$	B_n		$r_i = e_i$	1
A_n, $n \geq 2$	$G_2 = A_2$		r_1=(-1,1,0,...,0) r_2=(0,-1,1,0,...,0) ⋮ r_n=(0,...,0,-1,1) ($r_i \in R^{n+1}$)	$n+1$
D_n, $n \geq 4$	C_n $F_4 \cong D_4$		r_1=(-1,-1,0,...,0) r_2=(1,-1,0,...,0) ⋮ r_n=(0,...,0,1,-1) ($r_i \in R^n$)	4
E_8	–		r_1=(-1,1,0,...,0) r_2=(0,-1,1,0,...,0) ⋮ r_7=(0,...,0,-1,1) r_8=($\frac{1}{2}$, $\frac{1}{2}$,$\frac{1}{2}$,$\frac{1}{2}$,$\frac{1}{2}$,-$\frac{1}{2}$,-$\frac{1}{2}$,-$\frac{1}{2}$) ($r_i \in R^8$)	1
E_7	–		r_1=(-1,1,0,...,0) ⋮ r_6=(0,...,0,-1,1,0) r_7=($\frac{1}{2}$, $\frac{1}{2}$,$\frac{1}{2}$,$\frac{1}{2}$,-$\frac{1}{2}$,-$\frac{1}{2}$,-$\frac{1}{2}$,-$\frac{1}{2}$) ($r_i \in R^8$)	2
E_6	–		r_1=(0,-1,1,0,...,0) ⋮ r_5=(0,...,0,-1,1,0) r_6=($\frac{1}{2}$,$\frac{1}{2}$,$\frac{1}{2}$,$\frac{1}{2}$,-$\frac{1}{2}$,-$\frac{1}{2}$,-$\frac{1}{2}$,-$\frac{1}{2}$) ($r_i \in R^8$)	3

reasons that will soon become clear).

Theorem 2.6.2 (Witt's Theorem): *If $\Lambda = \Lambda_R$ (i.e. if Λ is generated by its root vectors), then Λ can be decomposed into a direct sum of the root lattices Z^n, A_n, D_n, E_n.*

(For a proof see WIT and KNE.)

This theorem provides us with a hint of the usefulness of root lattices which will become more graphic in the following pages. Most lattices known are either root lattices themselves, or constructed from them in some way (notable exceptions are Λ_{24}, O_{23}, O_{24}, Λ_{16} and K_{12}).

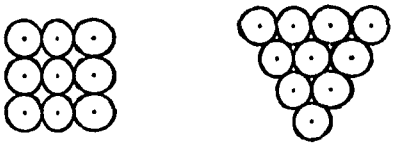
The root lattices and their duals solve many packing-type problems. (See Figure 3 for examples of these problems in 2 dimensions). The densest packing in 2-dimensions is A_2, as any billiard player can assure you. The best sphere packings known in dimensions 3-8 are A_3, D_4, D_5, E_6, E_7 and E_8 (the lattice points are the centers of the non-overlapping n-spheres). The "dual", so-to-speak, of the sphere packing problem is the Covering Problem: completely cover R^n with (overlapping) spheres of fixed radius in such a way as to minimize the overlap. A_2 is again the best in 2-dimensions, while A_k^*, for $k = 3, \ldots, 8$ are the best known in their respective dimensions. Another famous question is the kissing number problem: fix an n-sphere in R^n, and try to maximize the number of non-overlapping n-spheres of equal radius that just touch the central one. Isaac Newton and David Gregory apparently had a debate about this in 1694 — Newton thought that in 3-dimensions only 12 ball bearings can simultaneously "kiss" a central one, while Gregory thought 13 could be possible. Mathematics has since sided with Newton. It has been proven in 2, 3 and 8 dimensions that A_2, A_3 and E_8 are the best (in 8 dimensions, for example, 240 spheres can be arranged in this way), and it is thought that D_4, D_5, E_6 and E_7 are also the best.

The classical representation groups of A_n, B_n, C_n and D_n are SU(n+1), SO(2n+1), Sp(2n) and SO(2n), respectively. This will prove useful in translating from the language of (momentum) lattices to the language of (gauge) groups. More on this in the next section and in §1, Chapter 3.

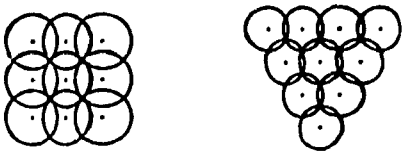
The root lattices are the solutions to another fundamental problem: enumerate all finite reflection groups, also called Coxeter groups. An irreducible finite reflection group (the reducible ones are simply direct products of the irreducible ones) can be defined as follows: engrave on the surface of a sphere

Figure 3: Sphere Packings in Two-Dimensions

3(a) The sphere packing problem



3(b) The thinnest covering problem



3(c) The kissing problem

a spherical simplex, all of whose dihedral angles are of the form $\frac{\pi}{p_i}$, and consider the group generated by reflections in its walls.

It was shown by Coxeter that this definition is equivalent to the following: a finite reflection group is a finite group with elements R_i satisfying

$$R_i^2 = 1 \text{ and } (R_i R_j)^{p_{ij}} = 1$$

for integers p_{ij}. (Here, $\frac{\pi}{p_{ij}}$ is the angle between the ith and jth walls, and R_i is the reflection in the ith wall.)

The crystallographic reflection groups have each p_{ij}=2, 3, 4 or 6. Associate with these groups a Coxeter diagram where each wall of the simplex (the interior of which is called the fundamental region of the group) is represented by a node, and nodes i and j are connected by 0, 1, 2 or 3 lines depending on whether p_{ij}=2, 3, 4 or 6, respectively. Then the Coxeter diagram is just a disjoint union of the diagrams of A_n, B_n, C_n, D_n, E_n, F_4 and G_2. (Non-crystallographic reflection groups have also been classified). From these we can get lattices by choosing vectors orthogonal to the reflecting hyperplanes.

We can get a reflection group out of a lattice, as well. Define a root vector for Λ to be a primitive $r \in \Lambda$ such that the reflection in the hyperplane orthogonal to r, given by $x \to x - \frac{2x \cdot r}{r \cdot r} r$, is a symmetry of Λ (r is said to be primitive if $\lambda r \in \Lambda \Longrightarrow \lambda \in Z$). These reflections generate a subgroup of $\mathrm{Aut}(\Lambda)$, sometimes called the Weyl group. If Λ is self-dual, these roots are precisely the vectors of norm 1 or 2 in Λ — hence in agreement with our earlier definition of root vector.

2.7 Gluing Theory

The concept of direct sum is an immensely valuable tool for the analysis of lattices, as Table 1 indicates. For example, it enables us to effectively ignore lattices containing unit vectors. But we can do even better than direct sum. A generalization of it is called gluing theory. With it we can breakdown even some indecomposable lattices, and can essentially do for norm 2 vectors what direct sums did to norm 1 vectors.

Consider some n-dimensional integral lattice Λ, and suppose the direct sum $\Lambda' = \Lambda_1 \oplus \cdots \oplus \Lambda_k$ is a sublattice of Λ also of dimension n. (Every lattice Λ has infinitely many such lattices Λ'. We are most interested, though, in cases where the Λ_i are all root lattices, which is somewhat rarer). Then any

$x \in \Lambda$ can be uniquely written as $x = x_1 + \cdots x_k$, where $x_i \in R\Lambda_i$. In fact, we can pin down these x_i even more (of course, in general $x_i \notin \Lambda$, since usually $\Lambda' \neq \Lambda$).

Choose any $y \in \Lambda_i$. Then $x_i \cdot y = x \cdot \hat{y} \in Z$, where

$$\hat{y} = \underbrace{0 \oplus \cdots \oplus 0}_{i-1} \oplus y \oplus \underbrace{0 \oplus \cdots \oplus 0}_{k-i} \in \Lambda.$$

Thus x_i must be in Λ_i^*.

Λ integral $\Longrightarrow$ Λ_i integral $\Longrightarrow$ Λ_i is a sublattice of Λ_i^*. Consider the quotient $G_i = \Lambda_i^*/\Lambda_i$, called the glue group of Λ_i. It's abelian, with order $|\Lambda_i|$ (see Theorem 1 below). The elements in G_i are sets (i.e. cosets) of vectors — choose a representative $g_j^{(i)}$ from each of these $|\Lambda_i|$ cosets. It's conventional to choose $g_j^{(i)}$ to have as small a norm as possible (so, for example, one of the $g_j^{(i)}$, corresponding to the j that labels the coset Λ_i, will be the zero vector), but there's a problem with this convention, as we will soon see. These $g_j^{(i)}$ are called the glue vectors of Λ_i.

Then $x \in \Lambda$ can be uniquely expressed as

$$x = y_1 + \cdots y_k + g,$$

where $y_i \in \Lambda_i$, and where g (also called a glue vector, unfortunately) is a sum

$$g = g_{j_1}^{(1)} + \cdots + g_{j_k}^{(k)}.$$

Define the set $G = \Lambda/\Lambda' = \{g = g_{j_1}^{(1)} + \cdots + g_{j_k}^{(k)} \mid \exists x \in \Lambda, y \in \Lambda',$ satisfying $x = y + g\}$. Then G is a subgroup of $G_1 \times \cdots \times G_k$, and is called (unfortunately) the glue group of the gluing of Λ by $\Lambda_1, \ldots, \Lambda_k$. We may write $\Lambda = (\Lambda_1, \ldots, \Lambda_k, G)$. G has order $(\prod_{i=1}^k |\Lambda_i|^{\frac{1}{2}})/|\Lambda|^{\frac{1}{2}}$ (see Theorem 1 below), and consists of all glue vectors of the gluing. G must satisfy two properties in order for Λ to be an integral lattice. First, it must be closed under addition, and secondly $g_1 \cdot g_2$ must be an integer for any $g_1, g_2 \in G$ (note that although the first property is always satisfied by $\prod G_i$, the second one is only if Λ is self-dual). The first condition is usually satisfied by specifying, instead of all of G, some generators $g_1, \ldots, g_\ell$ and defining G to be the Z-span of them. The second one is then verified by calculating each product $g_i \cdot g_j$ (column 3 of Table 3 is designed to make this task a little easier, at least when the components Λ_i are root lattices).

See Figure 4 for examples of 2-dimensional gluings (the origin is labelled o, while the glue vectors are denoted by bullets).

Direct sums arise when the glue group G consists only of the zero glue, and thus are a special case of gluings.

These values for $|G_i|$ and $|G|$ follow from:

Theorem 2.7.1 *Let Λ' be a sublattice of Λ such that $R\Lambda' = R\Lambda$. Then Λ/Λ' is an abelian group of order $\sqrt{\frac{|\Lambda'|}{|\Lambda|}}$. In particular, Λ^*/Λ has order $|\Lambda|$ for integral Λ.*

proof: Let U be the non-singular Z-matrix satisfying $M' = UM$, where M' and M are the generator matrices of Λ' and Λ, respectively. Then $A' = UAU^T$, so $|\Lambda'| = |U|^2|\Lambda|$, and it suffices to show that $|\Lambda/\Lambda'| = ||U||$.

Express U as the product $U = U_k \cdots U_1$ of elementary row matrices and define Λ_i for $i = 1, \ldots, k$ recursively as being the lattice whose generator matrix is $M_i = U_i M_{i-1}$, where $M_0 = M$. Then $\Lambda/\Lambda' = (\Lambda/\Lambda_1) \times (\Lambda_1/\Lambda_2) \times \cdots \times (\Lambda_{k-1}/\Lambda_k)$. Also, $|\Lambda_{i-1}/\Lambda_i|$ trivially equals $||U_i||$. Therefore $|\Lambda/\Lambda'| = |\Lambda/\Lambda_1| \times \cdots \times |\Lambda_{k-1}/\Lambda_k| = ||U_1|| \times \cdots \times ||U_k|| = ||U||$.

Example 2.7.1 *A simple example of a gluing is $\Lambda_i^* = \langle \Lambda_1, \ldots, \Lambda_k, G_1, \ldots, G_k \rangle$.*

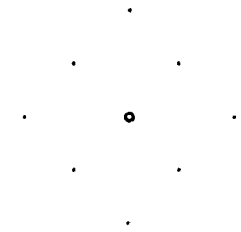
It is standard practice to represent the glue vector $g_j^{(i)}$ of the component Λ_i by $[j]$, and the glue vector $g = g_{j_1}^{(1)} + \cdots + g_{j_k}^{(k)}$ of Λ by $[j_1 \ldots j_k]$.

Call an integral lattice *saturated* if the sublattice Λ_R spanned by the vectors of norm 1 and 2 in Λ is n-dimensional (these vectors are called root vectors, and Λ_R is called the root lattice of Λ). Witt's Theorem (see the previous section) tells us that Λ_R is a direct sum of the indecomposable root lattices Z, A_n, D_n and E_n. Thus any saturated lattice can be formed by gluing together various root lattices. For this reason, the components Λ_i in a gluing are almost always chosen to be the root lattices.

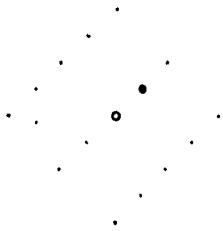
Not all lattices are saturated — some examples are $2Z$ and Λ_{24} — but self-dual lattices, at least in the smaller dimensions, often are. The first unsaturated self-dual lattice is 19-dimensional. 26 of the 27 Type II lattices

Figure 4: The Gluings of D_2

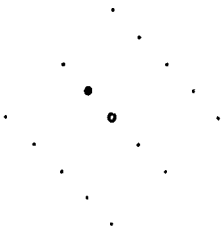
$D_2 = D_2[0]$



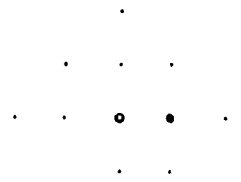
$D_2[1]$



$= D_2[3]$



$D_2[2] = Z^2$



of dimension $n \leq 24$ are saturated (the Leech lattice Λ_{24} is the sole exception). Besides Λ_{24}, two other important unsaturated self-dual lattices are the shorter and odd Leech lattices O_{23} and O_{24}.

One of the greatest triumphs of the gluing method was in Niemeier's enumeration of the 24 24-dimensional Type II lattices (see Table 4).[2] The remainder of this section will be devoted to studying the glue groups of the root lattices. The relevant data is collected in Table 3 (which should be read in conjunction with Table 2). The glue vectors in the Dynkin diagrams are labelled by bullets; for convenience we have written, e.g.,

$$(\{\frac{i}{n+1}\}^j, \{\frac{-j}{n+1}\}^i) \text{ for } (\underbrace{\frac{i}{n+1}, \ldots, \frac{i}{n+1}}_{j}, \underbrace{\frac{-j}{n+1}, \ldots, \frac{-j}{n+1}}_{i}).$$

Also included in Table 3 are the dot products of the glue vectors among themselves (for D_n and E_6 these have been arranged in Gram-like fashion), as well as the glue groups (note that the glue group of D_n depends on whether n is even or odd). Note that $[0]=(0,\ldots,0)$ is always a glue vector, and that the number of glue vectors of A_n, for example, is $|A_n| = n+1$ (so, for example, the self-dual root lattices Z^n and E_8 have only [0] for a glue vector, and hence enter as components into gluings only trivially, through direct sums).

Example 2.7.2 *$D_{n+m} = \langle D_n, D_m, [20], [02] \rangle$ for any $m,n \geq 4$.*

Example 2.7.3 *One of the most important gluings is $D_n^+ \equiv D_n[1] \equiv \langle D_n, [1] \rangle = D_n \cup (D_n + (\frac{1}{2}, \ldots, \frac{1}{2}))$. D_n^+ is integral iff n is a multiple of 4 (this should be obvious), in which case it is self-dual (this will be shown in a couple of different ways in the following sections). D_n^+ is Type I iff $n \equiv 4 \pmod 8$, and Type II iff $n \equiv 0 \pmod 8$. $D_4^+ = Z^4$, and $D_8^+ = E_8$. $|D_n^+|$ is 1 for n even, $\frac{1}{4}$ for n odd. D_n^+ is congruent to $D_n[3]$, but $D_n[2] = Z^n$ so equals D_n^+ iff $n=4$ (See Figure 4 for D_2 and $D_2[i]$). [1] is the deep hole for D_n (for $n \geq 1$; for $n=2$, [2] is the deep hole — see Figure 4), so geometrically D_n^+ consists of two copies of D_n superimposed in such a way that each copy plugs up a number of deep holes of the other. (Incidently, it turns out that every non-self-dual root lattice has only one type of deep hole, and one of its glue vectors points at that deep hole. This is not generally the case for other lattices.)*

[2] Table 4 was based on Table 16.1 of CS

Table 3: The Non-zero Glue Vectors

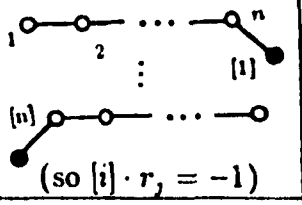
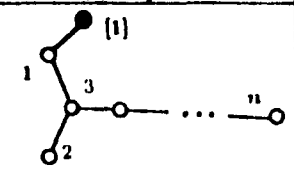
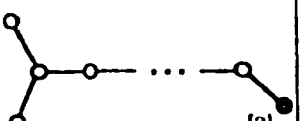
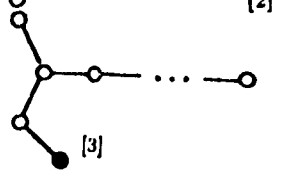
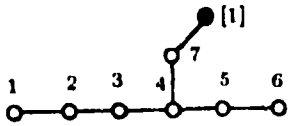
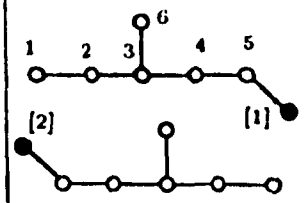
	Glue Vectors	$[i]\cdot[k]$	Dynkin Diagrams	Reps	Glue Gp
A_n	$[i]=$ $(\{\frac{i}{n+1}\}^j, \{\frac{-i}{n+1}\}^i)$ for $i=1,\dots,n$ $(j=n+1-i)$	For $i\geq k$, $[i]\cdot[k]=\frac{jk}{n+1}$	 (so $[i]\cdot r_j=-1$)	ten-sor of rank i	C_{n+1} ($[i]+[k]=[i+k]$)
D_n	$[1]=(\{\frac{1}{2}\}^n)$			s	$C_2\times C_2$ (for n even)
	$[2]=(\{0\}^{n-1},1)$	$\begin{matrix}\frac{n}{4} & \frac{1}{2} & \frac{n-2}{4}\\ \frac{1}{2} & 1 & \frac{1}{2}\\ \frac{n-2}{4} & \frac{1}{2} & \frac{n}{4}\end{matrix}$		v	C_4 (for n odd)
	$[3]=(-\frac{1}{2},\{\frac{1}{2}\}^{n-1})$			cs	($[1]+[2]=[3]$)
E_7	$[1]=$ $(\{\frac{-1}{4}\}^7, \frac{7}{4})$	$\frac{7}{2}$		56	C_2
E_6	$[1]=$ $(-\frac{1}{2},\frac{5}{4},\{-\frac{1}{6}\}^5,\frac{1}{2})$ $[2]=$ $(-\frac{1}{2},\{\frac{1}{6}\}^5,-\frac{5}{6},\frac{1}{2})$	$\begin{pmatrix}\frac{4}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{4}{3}\end{pmatrix}$		27 27*	C_3

Table 4: The 24-dimensional Type II (Niemeier) Lattices

Components	Glue Vector Generators g_i	Total # of Glue Vectors	$\lvert G_1\rvert$	$\lvert G_2\rvert$
D_{24}	[1]	2	1	1
$D_{16}E_8$	[10]	2	1	1
E_8^3	[000]	1	1	6
A_{24}	[5]	5	2	1
D_{12}^2	[(12)]	4	1	2
$A_{17}E_7$	[31]	6	2	1
$D_{10}E_7^2$	[110], [301]	4	1	2
$A_{15}D_9$	[21]	8	2	1
D_8^3	[(122)]	8	1	6
A_{12}^2	[15]	13	2	2
$A_{11}D_7E_6$	[111]	12	2	1
E_6^4	[1(012)]	9	2	24
$A_9^2D_6$	[240], [501], [053]	20	2	2
D_6^4	[even perms of {0123}]	16	1	24
A_8^3	[(114)]	27	2	6
$A_7^2D_5^2$	[1112], [1721]	32	2	4
A_6^4	[1(216)]	49	2	12
$A_5^4D_4$	[2(024)0], [33001], [30302], [30033]	72	2	24
D_4^6	[111111], [0(02332)]	64	3	720
A_4^6	[1(01441)]	125	2	120
A_3^8	[3(2001011)]	256	2	1344
A_2^{12}	[2(11211122212)]	729	2	95 040
A_1^{24}	[1(000001010011001101011111)]	4096	1	244 823 040
Λ_{24}	–	–	1	1

It often happens that $\mathrm{Aut}(\Lambda)$ should be computed (e.g. when using the Minkowski-Siegal mass formulae). When all the components are root lattices this is relatively easy to do. Let G' be the direct product of the Weyl groups of the components (the Weyl group is generated by the reflections in the simple roots, and is a normal subgroup with low index in the group of all automorphisms of the root lattice). Let G'' be the permutation group resulting from the automorphisms of Λ which map each component onto itself and which permute each component's glue vectors. It is a subgroup of the direct product of the symmetry groups of the Dynkin diagrams of the components (for example, the symmetry group of the Dynkin diagram for D_n , $n > 4$, consists of only two elements: 'parity', which flips [1] and [3], and the identity. For $n = 4$ there are the additional 'triality' conjugations). And finally, let G''' be the group of all permutations of the components that arise from $\mathrm{Aut}(\Lambda)$ (this is a subgroup of $S_{k_1} \times \cdots \times S_{k_\ell}$ if there are exactly k_i components equal to Λ_i, where $\Lambda_i \neq \Lambda_j$ whenever $1 \leq i < j \leq \ell$). Then $\mathrm{Aut}(\Lambda) = G' \times G'' \times G'''$. In particular, if G'' and G''' are as big as possible we will only get out the automorphism group of the direct sum of the components, so the automorphism group of a gluing is always a subgroup of that for the corresponding direct sum. For example, G'' and G''' always have only one element for D_n^+, so its automorphism group is half as large as that for D_n ($n > 4$).

What makes this decomposition of $\mathrm{Aut}(\Lambda)$ so useful is that G''' and G'''' in practice tend to be small and easy to calculate (see Table 4 for their orders for the Niemeier gluings), and the Weyl groups (and hence G'') are all known.

My glue vectors in Table 3 differ from those given elsewhere, notably CS. Unlike them I needed the Dynkin diagrams shown in Table 3 to be as simple as possible, for a number of reasons, something Conway and Sloane were apparently indifferent about. They were predominantly interested in choosing vectors of minimal length — e.g. their [1] for E_7 had norm 3/2, while mine has norm 7/2. Also, I adopted their coordinatization of the simple roots, unlike many people (see, e.g. LAM4), so my glue vectors reflected this.

I mentioned last section that the dual of a root lattice was its weight lattice. We have of course $g_i \in \Lambda_i^*$ — in fact it turns out that the glue vectors are weight vectors. This is how gluings will be found to relate to representations: in short the root lattice tells us the Lie algebra and the corresponding adjoint Lie group, and the glue vectors give us the representations (or the relevant covering) of that Lie group.

We will start this discussion with a few definitions.

Definition 2.7.1 *A Lie algebra L is a vector space furnished with an antisymmetric bilinear product* $[-,-]$ *satisfying the Jacobi identity* $[[a,b],c]+[[b,c],a]+[[c,a],b]=0$ *(we shall assume the Lie algebra is finite dimensional).*

A Lie algebra is often characterized in physics by its structure constants. Choose any basis b_i for L. Since L must be closed under the bilinear product, we can write $[b_j, b_k] = c_i^{jk} b_i$ for some scalars c_i^{jk} (called the structure constants of L relative to b_i).

Definition 2.7.2 *Let V be a vector space, and L be a Lie algebra. A mapping* ρ *which assigns to each* $a \in L$ *a linear transformation* $\rho(a)$ *on V is called a* representation *of L on V if it satisfies:*

1. $\rho(\lambda a + \mu b) = \lambda\rho(a) + \mu\rho(b)$ *for all* $a, b \in L$, $\mu, \lambda \in R$;

2. $\rho([a,b]) = \rho(a)\rho(b) - \rho(b)\rho(a)$.

For example, the trivial representation sends each vector in L to the 0 transformation. A more important example is the adjoint representation rho_A, for which $\rho_A(b_j)$ is the matrix whose ik-th entry is c_i^{jk}. The adjoint representation has dimension equal to the rank of the algebra (the dimension of a representation is the dimension of the vector space V on which it acts).

Let ρ be a representation of L on V. Then a linear functional ω on L is called a weight of ρ if there exists a nonzero $v \in$ V such that $\rho(a)(v) = \omega(a)v$ for each $a \in$L. v is called a weight vector belonging to the weight ω (in other words the weight vector is an eigenvector of each matrix $\rho(a)$, and $\omega(a)$ is the corresponding eigenvalue).

The weights of the adjoint representation are called roots, and its weight vectors are simply the root vectors of L that we are by now quite familiar with.

The representation ρ is called irreducible if no proper subspace W of V is invariant under ρ — i.e. if $\rho(a)$W$\subseteq$W for each $a \in$ L, then W = V or 0. The matrices in a reducible representation can be simultaneously put in block diagonal form, and so the reducible representation can be expressed as a direct sum of irreducible representations.

One of the weights, called the maximal, or highest, weight of ρ can be singled out; it turns out that an irreducible representation ρ is characterized

by its maximal weight ω_ρ (i.e. two irreducible representations are isomorphic iff their maximal weights are equal). Then there is a natural, one-to-one correspondence between the family of all n-tuples $(m_1, \ldots, m_n)$ of non-negative integers (where n is the rank of the Lie algebra — i.e. the number of simple roots, or the number of nodes on its Dynkin diagram) and the family of all (equivalence classes of) finite-dimensional irreducible representations of L. The m_i can be considered to be the values of the maximal weight on the simple roots.

Fundamental weights are those corresponding to the n-tuples with $n-1$ 0's and one 1. They (or rather their corresponding weight vectors) form a basis of the weight lattice dual to the basis of simple roots of the root lattice. The irreducible representations which have as their maximal weight a fundamental weight, are called fundamental representations. All representations can be obtained from the fundamental ones by decomposing their tensor products. Our glue vectors (as opposed to those of Conway and Sloane) are fundamental weight vectors (or, actually, their negatives are), so to each of our glue vectors corresponds an irreducible representation (see Table 3). (The glue vector [0] always corresponds to the adjoint representation since its weight vectors are simply the root vectors.)

Closely related to this is that, up to isomorphism, there are only finitely many complex semisimple Lie groups (i.e. connected Lie groups whose Lie algebra is semisimple) with a given Lie algebra. The smallest is called the adjoint group, and has a trivial center. The others are covers of this group, and the largest is simply-connected For example, there are two semisimple Lie groups corresponding to the Lie algebra A_1. One of these is PGL(2)=SL(2)/Z_2, and the other is SL(2).

A less trivial example is D_n. SO($2n$), the group of rotations in R^{2n}, is the classical Lie group corresponding to the Lie algebra D_n. It has a trivial center. Its simply-connected covering group is the spinor group Spin($2n$). They have the same Lie algebra, and hence the same root lattice. The representations of Spin($2n$) fall into 4 conjugacy classes: adjoint or scalar (corresponding to [0]), vector ([2]), and the two irreducible spinor representations of positive and negative chirality (spinor [1] and conjugate spinor [3]). D_n^+, at least for even n, corresponds to the double cover Spin($2n$)/Z_2 of SO($2n$) because we want only half of the conjugacy classes of Spin($2n$) (namely the adjoint and spinor representations).

2.8 Lamination and the Leech Lattice

Lamination is one of three standard ways of constructing higher dimensional lattices from smaller ones. Its primary significance, at least for our purposes, lies in its ability to provide a construction/definition of the Leech lattice, essentially by starting from nothing. Laminated lattices are the densest packings -- lattice and non-lattice — known in each dimension ≤ 29, except for dimensions 10-13 (the Coxeter-Todd lattice K_{12} holds the record for 12 dimensions, and non-lattice packings are densest in dimensions 10, 11, and 13. Denser lattices are also known for all dimensions ≥ 30.)

Laminated lattices are defined to be the densest "layered" lattices of each dimension, and are listed in Figure 5[3] (unfortunately, I was unable to label two of the lattices in Figure 5: Λ_{12}^{Mid} and Λ_{13}^{Mid}). For dimensions ≤ 8, these are similar (but not congruent) to root lattices. In particular, letting Λ_n denote a laminated lattice of dimension n, we have:

$$\Lambda_1 \cong Z \cong A_1 \ , \ \Lambda_2 \cong A_2 \ , \ \Lambda_3 \cong A_3 = D_3 \ , \ \Lambda_4 \cong D_4$$

$$\Lambda_5 \cong D_5 \ , \ \Lambda_6 \cong E_6 \ , \ \Lambda_7 \cong E_7 \ , \ \Lambda_8 \cong E_8.$$

(These similarities were established by Leech in 1969.) Other laminated lattices include the Leech lattice Λ_{24} and the Barnes-Wall lattice Λ_{16}.

Λ_n is defined recursively as follows:

Let $\Lambda_1 = 2Z$. At the nth step stack copies of Λ_{n-1} as densely as possible, while preserving the minimal norm of 4 — the resulting lattice is Λ_n.

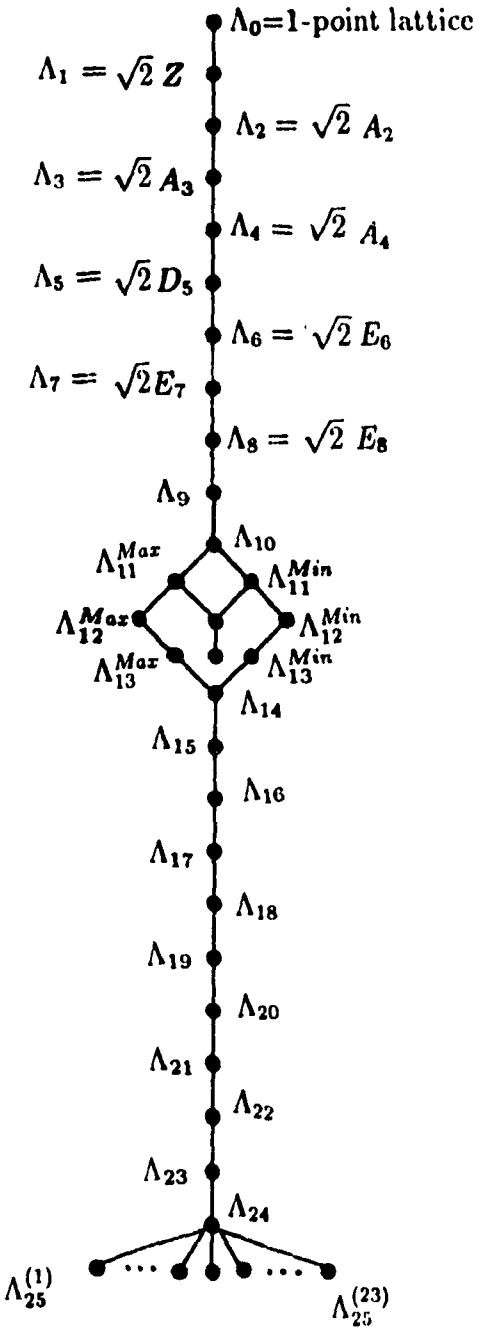
Equivalently, consider all n-dimensional lattices containing at least one sublattice Λ_{n-1}, whose minimal norm is 4. Λ_n are simply those of smallest determinant.

Sometimes the recursion is taken to start with the 1-point lattice Λ_0, rather than with Λ_1. Since Λ_0 is the only 0-dimensional lattice, or since Λ_1 is the only 1-dimensional lattice with minimal norm μ=4, lamination provides a very appealing definition for the Leech lattice, the unique laminated lattice of 24 dimensions.

Obviously the specific normalization chosen is not very important. The reason $\Lambda_1 = 2Z$ was chosen, rather than, for example, $\Lambda_1 = Z$, was so that Λ_{24} would be both integral and self-dual. With this normalization, Λ_n is integral iff $n \leq 24$. The choice of normalization becomes important only

[3] Figure 5 is based on Figure 6.1 in CS

Figure 5: The Laminated Lattices



when we demand each Λ_n be the lattice with the smallest determinant among all *integral* lattices having the given minimal norm and having some Λ_{n-1} as a sublattice. The shorter Leech lattice O_{23} is the unique integral laminated lattice of norm 3 in 23 dimensions, but $\Lambda_{23} \neq \sqrt{\frac{4}{3}} O_{23}$.

In general there isn't a unique Λ_n for a given n (see Figure 5), although Λ_n is unique for $n \leq 10$, and $14 \leq n \leq 24$. All Λ_n are known for $n \leq 25$, and at least one is known for $26 \leq n \leq 48$. There are exactly 23 different Λ_{25} (one for each type of deep hole in Λ_{24}), and there is good reason to believe that the number of different Λ_{26} is well over 75 000.

Of course, Λ_{24} is a sublattice of each Λ_n, $n \geq 24$. It has been shown that although not every Λ_n, for $n \leq 24$, is a sublattice of some Λ_{n+1}, every such Λ_n is a sublattice of Λ_{24}. For example, Λ_{20} is obtained from Λ_{24} by equating any 4 coordinates to 0.

Each lattice Λ_{n+1} is the union of translations of some lattice Λ_n. A simple geometric picture of how Λ_{n+1} is obtained from Λ_n can be found. Λ_{n+1} arises when you place each layer (roughly speaking, each copy of Λ_n) in such a way that each point in the layer is directly opposite a deep hole of the adjacent layer. This structure can be clearly seen in the case of $\Lambda_2 \cong A_2$ in Figure 1. When there are more deep holes than points in the layers, there are several possibilities for these layerings. Many of these possibilities will yield non-lattices, but it seems that always at least one is a lattice.

Usually, each of the parallel layers of Λ_{n+1} is just a copy of some Λ_n. But in dimensions 8 and 24, for example, "density doubling" occurs as adjacent layers can actually be merged. For instance, it is possible to fit together 2 copies of D_8, preserving the minimal norm. The result is $D_8^+ = E_8 \cong \Lambda_8$. A similar situation applies to the Leech lattice. This explains in part the remarkable properties possessed by E_8 and Λ_{24}.

There are several alternatives for defining Λ_{24}, but the way given above is perhaps the most appealing. Its generator matrix is given in Figure 6 (for readability, only its nonzero entries are shown). Λ_{24} is even and self-dual, and is the only such lattice in 24 dimensions which cannot be obtained by gluing root lattices. In fact, Λ_{24} contains no roots — its minimal norm $\mu = 4$, not 2. This is a very distinguishing property, for the only self-dual lattices of dimension ≤ 24 containing no roots are the Leech lattice Λ_{24}, the 23-dimensional shorter Leech lattice O_{23}, and the 24-dimensional odd Leech lattice O_{24}. O_{23} and O_{24} both have $\mu = 3$, and are intimately associated

Figure 6: The Leech Lattice Generator Matrix

$$\frac{1}{\sqrt{8}}\begin{pmatrix}
8 & & & & & & & & & & & & & & & & & & & & & & & \\
4 & 4 & & & & & & & & & & & & & & & & & & & & & & \\
4 & & 4 & & & & & & & & & & & & & & & & & & & & & \\
4 & & & 4 & & & & & & & & & & & & & & & & & & & & \\
4 & & & & 4 & & & & & & & & & & & & & & & & & & & \\
4 & & & & & 4 & & & & & & & & & & & & & & & & & & \\
4 & & & & & & 4 & & & & & & & & & & & & & & & & & \\
2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & & & & & & & & & & & & & & & & \\
4 & & & & & & & & 4 & & & & & & & & & & & & & & & \\
4 & & & & & & & & & 4 & & & & & & & & & & & & & & \\
4 & & & & & & & & & & 4 & & & & & & & & & & & & & \\
2 & 2 & 2 & 2 & & & & & 2 & 2 & 2 & 2 & & & & & & & & & & & & \\
4 & & & & & & & & & & & & 4 & & & & & & & & & & & \\
2 & 2 & & & 2 & 2 & & & 2 & 2 & & & 2 & 2 & & & & & & & & & & \\
2 & & 2 & & 2 & & 2 & & 2 & & 2 & & 2 & & 2 & & & & & & & & & \\
2 & & & 2 & 2 & & & 2 & 2 & & & 2 & 2 & & & 2 & & & & & & & & \\
4 & & & & & & & & & & & & & & & & 4 & & & & & & & \\
2 & & 2 & & 2 & & & 2 & 2 & 2 & & & & & & & 2 & 2 & & & & & & \\
2 & & & 2 & 2 & 2 & & & 2 & & 2 & & & & & & 2 & & 2 & & & & & \\
2 & 2 & & & 2 & & 2 & & 2 & & & 2 & & & & & 2 & & & 2 & & & & \\
& 2 & 2 & 2 & 2 & & & & 2 & & & & 2 & & & & 2 & & & & 2 & & & \\
& & & & & & & & 2 & 2 & & & 2 & 2 & & & 2 & 2 & & & 2 & 2 & & \\
& & & & & & & & 2 & & 2 & & 2 & & 2 & & 2 & & 2 & & 2 & & 2 & \\
-3 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}$$

with Λ_{24}. In addition, Λ_{24} is the only self-dual lattice of under 32 dimensions with $\mu \geq 4$. (On the other hand, there are at least two 32-dimensional Type II lattices with $\mu = 4$, and at least two 48-dimensional Type II lattices with $\mu = 6$. And if we drop the assumption of self-duality, we get further nontrivial examples: both K_{12} and Λ_{16} — see Figure 2 — have $\mu = 4$ (the trivial examples are lattices scaled in unusual ways, like $15Z$)). Considerably more common are self-dual lattices which, though containing roots, don't contain enough to be saturated — i.e. which cannot be expressed as the gluings of root lattices. The first examples are 19-dimensional, and in higher dimensions become the norm rather than the exception.

Every root lattice has only 1 or 2 types of deep holes. K_{12} and Λ_{16} both have only 1. But Λ_{24} has an incredible 23. This is also the number of Niemeier lattices, and indeed there is a fairly natural correspondence between them. Related to this are Conway's 23 "Holy Constructions" of Λ_{24}. (These are given and elegantly proved by Borcherds in BOR.)

There are some interesting connections between Λ_{24} and the Lorentzian lattice $II_{25,1}$. For example, you can speak of fundamental (or simple) roots for indefinite lattices just as was done for Euclidean lattices. $II_{1,1}$ has 10 of them, $II_{17,1}$ has 19, and $II_{25,1}$ has infinitely many. The Weyl vector of a lattice has inner product -1 with all the fundamental roots; it exists for $II_{n,1}$ for n =9, 17 and 25, but not for $n = 33, 41, \ldots$. The Weyl vector for $II_{25,1}$ is $w = (0, 1, 2, \ldots, 24|70)$ — surprisingly it is a null vector, which means $\Lambda_w = (w^\perp \cap II_{25,1})/\langle w \rangle$ is a 24-dimensional Type II Euclidean lattice. Conway has shown in fact that $\Lambda_w = \Lambda_{24}$. In addition, he found that the Coxeter-Dynkin diagram of $II_{25,1}$ can be identified with Λ_{24} (the diagram has 1 node for each Leech lattice vector).

Λ_{24} was discovered in 1965 by John Leech in the context of sphere packing. After being introduced to it by John McKay (a former professor at McGill), Conway soon elevated it to its current position as a truly significant mathematical construct. It has been proved to have the highest kissing number of any 24-dimensional packing, and holds the record at the present time for the densest packing, thinnest covering, and best quantizer (see §6 for descriptions of these famous problems) in 24 dimensions. It has connections with Lie algebras that aren't yet fully understood. But it is perhaps in the theory of finite simple groups where its usefulness is most graphic.

Let $\cdot\infty$ be the infinite automorphism group of Λ_{24} — i.e. all isometries of R^{24} which are symmetries of Λ_{24}. $\cdot\infty$ contains translations, for example.

Let ·0 (pronounced "dotto") be the stabilizer of the origin — i.e. those transformations in ·∞ that fix the origin. It is what is usually meant by $\mathrm{Aut}(\Lambda_{24})$. It turns out only rotations are in ·0 — Λ_{24} is the only Type II lattice in under 32 dimensions which isn't invariant under any reflection (this is in contrast to the root lattices, whose reflections almost generate all automorphisms, and is intimately associated, as we learned in §6, with the fact that Λ_{24} has no roots). ·0 has about 8×10^{18} elements, which is about average for 24-dimensional Type II lattices (the automorphism group of D_{24}^{+} has order $\approx 2.5\times10^{7}$, while that corresponding to A_3^8 has order $\approx 4.5\times10^{23}$).

Like all lattices, Λ_{24} is invariant under $x \to -x$. Define $\cdot1 = \cdot0/Z$, where $Z = \{-I, I\}$ is the center of ·0 (so ·0 is a double covering of ·1). In addition, choose any vectors $x, y \in \Lambda_{24}$ of norm 4 and 6 respectively, and define ·2 and ·3 to be those subgroups of ·0 fixing x and y respectively (·2 and ·3 are independent of the choice of these vectors, since ·0 is transitive on both $\Lambda(2) = \{v \in \Lambda_{24} \mid v^2 = 4\}$ and $\Lambda(3) = \{v \in \Lambda_{24} \mid v^2 = 6\}$). It can be shown that ·1, ·2 and ·3 are all simple -- in fact they're the 5th, 11th and 12th largest sporadic finite simple groups. They were discovered by Conway in 1968.

A total of 12 sporadic groups are contained in ·1 (or ·0 or ·∞) as subgroups. Many of them are stabilizers of a "simplex" in Λ_{24}, just like ·1, ·2 and ·3 (the simplices in these cases are the point 0, and the segments with head and tail 0 and x, and 0 and y respectively). Daniel Gorenstein, himself a major finite simple group theorist, wrote:

> If Conway had studied the Leech lattice some 5 years earlier, he would have discovered a total of 7 new simple groups! Unfortunately he had to settle for 3. However, as consolation, his paper on ·0 will stand as one of the most elegant achievements of mathematics.[4]

On the negative side, Conway's work distracted many people into suspecting more sporadic simple groups could be found in the automorphism groups of other lattices. This turned out to be a dead end.

Also, the Leech lattice was used by R. L. Griess in 1980 to construct the Monster group, the largest of the 26 sporadic groups with almost 8.1×10^{53} elements. The construction, done by hand, involved calculations that have

[4] GOR, p 125

been called "truly prodigious" and "horrendous". Conway later produced a simplified construction also heavily using Λ_{24} (and with it a vector space of dimension 196 883). The Monster 'involves' 20 — and possibly 21 — of the sporadic groups, including the Conway groups ·1, ·2 and ·3 (G 'involves' H means that H is a quotient of a subgroup of G).

The Monster already has many uses. For example, in string theory it has been used (by George Chapline — see CHA) along with ·1 to construct a 26-dimensional anomaly-free superstring theory unifying an $E_8 \times E_8$ gauge invariance and O(8) supersymmetry. ·1 plays the role of the symmetry which combines these, while the representations of the Monster are the fermion states. Perhaps anticipating a theory like this, Freeman Dyson recently wrote:

> I have to confess to you that I have a sneaking hope, a hope unsupported by any facts or evidence, that sometime in the 21st century physicists will stumble upon the monster group, built in some unsuspected way into the structure of the universe. This is of course only a wild speculation, almost certainly wrong. The only argument I can produce in its favor is a theological one. We have strong evidence that the creator of the universe loves symmetry, and if he loves symmetry, what lovelier symmetry could he find than the symmetry of the Monster?[5]

[5] DYS, p. 53

3 The Self-Duality of Gluings

3.1 Statement and Motivations for the Problem

As mentioned earlier, all self-dual lattices of dimensions $n \leq 25$ have been found. For higher n (and in particular for the even lattices of dimension 24), these were obtained for the most part by gluing together root lattices. It would be interesting to see how Niemeier, for example, established the self-duality of his Type II lattices. That the gluings are integral isn't hard to check: the dot products between any two generator glue vectors g_i must all be integers. The theoretical and practical difficulty seems to be in determining whether the gluing yields a lattice with determinant 1.

Unfortunately Niemeier's work (see NIE) has not yet been translated into English (to my knowledge). The "modern" proof, due to B. Venkov (see CS, pp. 427-438) is indirect, translating the task into a problem of coding theory (namely, classify the self-dual codes over all rings Z_n and the field F_4), so his methods are unable to provide insight into our question (except possibly to suggest that the question is no longer relevant). However, in superstrings if not in mathematics there is reason for developing methods to calculate the determinant of a gluing. In the following sections I'll elaborate on 3 such methods, as well as describe a clever graphical technique, due to Lam, of computing determinants of certain $n \times n$ matrices. In this section I'll try to describe why superstrings would bother with such a question.

Consider the problem discussed in §5 of Chapter 1 of constructing heterotic strings from free boson or fermion fields. When we bosonize the fermions we find that the allowed "momenta" p form a set $\Lambda + t$, where t is a vector pointing in the "fermionic direction", and where Λ is an indefinite integral lattice in R^{N_L,N_R}, for $N_L = 24 - d$ and $N_R = 12 - d$ ($d+2$ is the number of space-time dimensions; presumably we are ultimately interested in $d = 2$). For some peculiar reason we write, for example, $p = (p_R|p_L)$ (so $p^2 = p_L^2 - p_R^2$). In addition, modular invariance of the partition function strongly suggests Λ be self-dual (see §4 of the next chapter).

Let $G = G_R \times G_L$ be the local symmetry group — it must be rank $N = N_R + N_L$, since our lattice has that as its dimension. These momenta are the weight vectors of the representation of G corresponding to the relevant particle. Now d of the N_R dimensions will correspond to space-time (since we're in the light cone gauge). Special relativity demands that the Lorentz

group SO(d)=$D_{\frac{d}{2}}$ be the symmetry of those d dimensions, so $G_R = D_{\frac{d}{2}} \times G'_R$ (the actual Lorentz group is SO(d+1,1), but in the light cone gauge this becomes SO(d) — only the transverse directions survive). If we let p_0 be the momentum components in the space-time directions (p_0 has $d/2$ components because SO(d) is of rank $d/2$), we get that p_0 must be a weight vector of $D_{\frac{d}{2}}$ — it will correspond to the adjoint (scalar or [0]) or fundamental (vector or [2]) representation if the particle is a boson, or to the spinor representations [1] or [3] if a fermion. Similar arguments will apply to the remaining symmetries $G'_R \times G_L$.

In general we won't know all of G, and we won't know all of the representations (i.e. the glue vectors). We will be given the Lorentz group $D_{\frac{d}{2}}$, certainly, and a gauge group G_g of rank N_g. We will be given a set of low mass particles (e.g. gravitons, gauge particles, and other particles, like quarks) whose momenta correspond to representations of G_g. This amounts to being given the root lattice Λ_0 (representing the momenta of the gauge particles) for $D_{\frac{d}{2}} \times G_g$, and a set of glue vectors g_i (representing the quarks, etc.). We must find a self-dual lattice Λ of dimension N in which can be found the $N_g + d/2$-dimensional lattice $\Lambda' = \langle \Lambda_0, g_i \rangle$.

Unfortunately $N_g + d/2$ can be much smaller than N. The most important special case involves the Standard Model (see §2, Chapter 1), where $d = 2$ and G_g =SU(3)×SU(2)×U(1), for a total rank of 5, while $N = 32$. In addition, we have (at present) experimental access to the low mass particles only (according to the theory they will in fact be massless, at least until they acquire a small mass via symmetry breaking), since the massive particles will have masses on the order of the Planck mass $M_P \approx 10^{19}$ Gev, or about the size of a bacterium. This means that we can only guess at the glue vectors corresponding to the massive particles. (Incidently, the massless states correspond to $p = (p_R|p_L)$ satisfying $p_R^2 \leq 1$, $p_L^2 \leq 2$.)

Suppose we guess at a rank N group G (of course one containing $D_{\frac{d}{2}} \times G_g$). Then before we add more glue vectors we will have an N-dimensional lattice $\Lambda = \langle \Lambda_G, g_1, \ldots, g_\ell \rangle$, where Λ_G is the root lattice of G. First of all, this lattice must be integral, otherwise our task is hopeless. This is trivial to verify (just check that the $\binom{\ell}{2}$ products $g_i \cdot g_j$ are in Z). Next, calculate its determinant. If it's 1, then we're done: we will have successfully constructed a string theory with the desired low energy symmetries and spectrum. Otherwise it will be necessary to add additional glue vectors g'_i to the gluing. This will have the

effect of lowering $|\Lambda|$, perhaps to 1.

This discussion, though necessarily incomplete, hopefully gives some idea of the motivation in superstrings behind calculating the determinant of gluings (see LAM4 for a more complete account). Lam has devised 2 methods of checking for self-duality: one involves manipulating theta series (I won't address that approach; it can be found in LAM4), and the other involves computing greatest common divisors of certain determinants (I'll discuss this in §3). In addition I have found 2 other techniques, which I will discuss in §2 and at the beginning of §3.

The mathematical problem can be stated in the following way.

Let $\Lambda_i, i = 1, \ldots, k$ be n_i-dimensional root lattices embedded in R^{m_i}, and define $n = n_1 + \cdots + n_k, m = m_1 + \cdots + m_k$. Consider the gluing Λ of these lattices $\Lambda_1, \ldots, \Lambda_k$ via the glue vectors $g_1, \ldots, g_\ell$. For $i = 1, \ldots, k$, let $r_1^i, r_2^i, \ldots, r_{n_i}^i \in R^{m_i}$ be the simple roots of Λ_i; these constitute a basis for each Λ_i, and (provided we read for each r_j^i the vector in R^m with components $(0^{m_1+\ \ +m_{i-1}}, r_j^i, 0^{m_{i+1}+\cdots+m_k})$) all of these simple roots together form a basis for the sum $\Lambda_1 \oplus \cdots \oplus \Lambda_k$. Given that the n-dimensional lattice Λ is spanned, by definition, by the $n + \ell$ vectors in $B = \{r_j^i, g_1, \ldots, g_\ell\}$, the problem is to determine if it is self-dual. We shall solve this by explicitly calculating $|\Lambda|$.

First note, however, that $\frac{|\Lambda'|}{|\Lambda|}$ is the square of an integer, provided Λ' is an n-dimensional sublattice of Λ. The reason is that because $\Lambda' \subseteq \Lambda$, the basis of Λ spans that of Λ', so $M' = UM$ for some $n \times n$ Z-matrix U, which implies $A' = UAU^T$ and hence $|\Lambda'| = |A'| = |U|^2|\Lambda|$. And of course $|U| \in Z$. This gives us the following practical test:

Theorem 3.1.1 *A gluing of $\Lambda_1, \ldots, \Lambda_k$ can be self-dual only if $|\Lambda_1| \cdots |\Lambda_k|$ is the square of an integer.*

The converse unfortunately fails (e.g. no gluing of $2A_2$ can be self-dual). This test is particularly easy to perform given that the determinants of each root lattice is known (see Table 2).

Let $b_1, \ldots, b_n$ be an enumeration of the n root vectors r_j^i, and let $b_{n+i} = g_i$. Define $\Lambda_{\{i_1, \ldots, i_n\}}$ to be that sublattice of Λ obtained by taking the Z-span of $\{b_{i_1}, \ldots, b_{i_n}\}$. We will get at Λ by looking at these sublattices, because it is not immediately obvious how to directly compute $|\Lambda|$ (B is not a basis).

Normally, $|\Lambda|$ would involve calculating the determinant of an $n \times n$ Gram matrix — a potential problem for large n ($n = 32$ is an important special

case). But because root lattices are involved, and because we have chosen as their basis vectors simple roots, the calculation of $|\Lambda_{\{i_1,\ldots,i_n\}}|$ can be done quickly, using a graphical method given in §4. The question becomes: can a way be found to get $|\Lambda|$ from $|\Lambda_{\{i_1,\ldots,i_n\}}|$. Two of the three methods I'll discuss do exactly that, and seem quite successful in practice. But the most promising is the third, given at the beginning of §3.

Strictly speaking, Λ is indefinite and not Euclidean as we will be assuming in this chapter. G must decompose into the direct product $G_R \times G_L$ (because the right- and left-movers cannot mix), and hence the root lattice must likewise decompose into the direct sum of two Euclidean root lattices (albeit one is negative definite). I see no reason, though, why the glue vectors must respect this decomposition and have all their non-zero components either only in right-moving or only in left-moving directions. If this isn't the case, and it seems doubtful that it is, then Λ cannot be decomposed in this way, and our assumption would appear to be groundless. However, the methods discussed in this chapter should generalize quite naturally to the more realistic indefinite case. The examples considered on the following pages will all be Euclidean, both for reasons of simplicity and for the wealth of explicit Euclidean gluings that are available (see, for example, Table 4).

3.2 Finding a Basis

The most natural way to determine whether a lattice is self-dual is to check its determinant. The most natural way to do this is to find a basis for it, and to use this basis to find a Gram matrix for the lattice.

One way of finding a basis for Λ given the generating vectors b_i is by row-reducing (over Z) the matrix

$$M = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n+\ell,1} & \cdots & b_{n+\ell,m} \end{pmatrix}$$

into the equivalent upper triangular matrix

$$T = \begin{pmatrix} * & * & \cdots & & & \\ 0 & * & \cdots & & & \\ \vdots & 0 & \ddots & & & \\ & & 0 & * & * & \cdots \\ 0 & \cdots & & 0 & \cdots & 0 \\ \vdots & & & \vdots & & \vdots \end{pmatrix}.$$

(In general M isn't a Z-matrix, but for the gluings of root lattices M will always be rational; letting p be a common denominator for all the entries of M, and noting that $\frac{1}{p}Z$ is a principal ideal domain, we see that Prop. 2.11 on page 339 of HUN guarantees that this decomposition is possible.) The desired basis of Λ is the collection of non-zero rows of T. The reduction procedure itself is easily programmable. The only complication is that M is a large matrix (e.g. $n = 36 - 2d$), and the resulting Gram matrix would not necessarily be sparse, discouraging direct computation of its determinant. Thus an alternative to this procedure would be desirable, and in view of the complication caused by a non-sparse Gram matrix, what would be preferred would be a method which yielded basis vectors closely related to the b_i.

This suggests that we should try to choose basis vectors from the set B. Now, it is certainly true that if we were dealing with vector spaces, some subset of $\{b_i\}$ must be a basis for R-span$\{b_i\}$. For a lattice, however, this isn't necesarily the case (see the following examples). Call the set B *redundant* if one of its subsets was a basis. Ultimately, we shall provide a necessary and sufficient test for redundancy, and also determine all possible subsets of B, if any, that form the basis. It will turn out that most practical situations are redundant, or become redundant after trivial manipulations.

Before introducing the method, a graphical representation will be introduced to make the following discussion particularly tangible. It will also be exploited, as we will learn in §4, in the calculation of determinants.

To each Lie algebra is associated its Dynkin diagram. These can be constructed as follows: represent each of its simple roots with a node, and connect 2 such nodes with a solid line iff the dot product of their corresponding roots is non-zero (since we're dealing here only with products of simply laced algebras, only single, undirected lines apply). We can define this similarly

for lattices — the result is a graphical representation of its Gram matrix. For lattices we shall call this the Coxeter diagram.

For example, the Coxeter diagram of D_8^2 is given in Figure 7(a). The lattice $\langle D_8^2, [11], [22] \rangle$ has the diagram shown in Figure 7(b). We see quite clearly here the geometric effect of gluings: the glue vectors [11] and [22] link up the orthogonal component lattices D_8 and D_8. A few points must be made. We have adopted the convention that the simple roots be represented by hollow nodes, and the glue vectors by solid ones. Also, dot products of -1 between vectors have been, as before, represented by single solid lines; here dot products of +1 are represented by broken lines. Dot products other than ± 1 will be explicitly indicated, as will any norm different from 2. In general, the Coxeter diagrams will be useful only for root lattices and their gluings; the Coxeter diagram for the Leech lattice, for example, is an impenetrable maze.

But the main problem is that there are now 8+8+2 $(=n_1+n_2+\ell = n+\ell)$ nodes, and the dimension of the whole lattice is only 8+8 $(=n_1+n_2 = n)$. In Dynkin diagrams the rank of the Lie algebra equals the number of nodes; we would like the number of nodes in these Coxeter diagrams to equal the dimension of the corresponding lattice. In particular, we would like the nodes to represent basis vectors, and not merely the generator vectors b_i.

As was discussed above, it would be most convenient if we could choose our basis from among the vectors in B. This would amount to removing ℓ of the nodes. For example, it turns out that the generating vectors b_i of $\langle D_8^2, [11], [22] \rangle$ are redundant; its correct Coxeter diagram can be seen in Figure 7(c) (there are 7 other equally valid choices of node removal available here, as we shall see in our examples).

Linear independence of vectors $g_1, \ldots, g_\ell$ of course simply means

$$a_1 g_1 + \cdots + a_\ell g_\ell = 0 \Longrightarrow a_1 = \cdots = a_\ell = 0$$

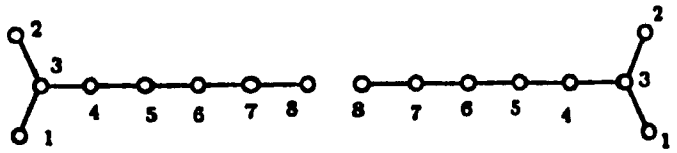
For our purposes we require something a little stronger.

Definition 3.2.1 *Say that $\{g_1, \ldots, g_\ell\}$ is Λ-independent if they are linearly independent, and if whenever we have integers n_i satisfying*

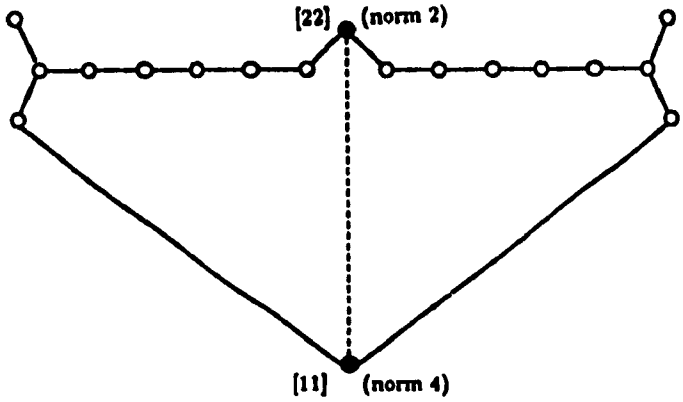
$$n_1 g_1 + \cdots + n_\ell g_\ell \in \Lambda$$

then for each i we have $n_i g_i \in \Lambda$.

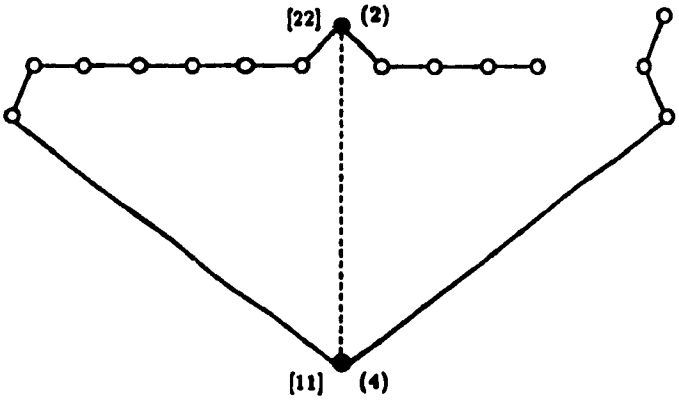
Figure 7: The Coxeter Diagrams of $\langle D_8^2, [11], [22] \rangle$



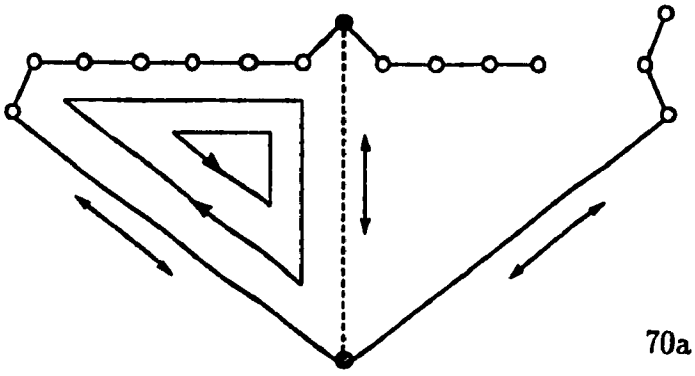
(a) The Coxeter diagram for D_8^2



(b) An (improper) Coxeter diagram for $\langle D_8^2, [11], [22] \rangle$



(c) A (proper) Coxeter diagram for $\langle D_8^2, [11], [22] \rangle$



(d) The 5 loops through [11]

We will be specifically interested in the cases where the glue vectors g_i for the sum $\Lambda' = \Lambda_1 \oplus \cdots \oplus \Lambda_k$ are Λ'-independent. Later we will discuss just how comprehensive our method can be, given this constraint of the glue vectors being Λ'-independent. An immediate consequence of this constraint is that the ℓ nodes to be removed from the set of $n + \ell$ generators b_i must all be root vectors. The actual motivation for demanding Λ'-independence is both more subtle and more significant, and will be shortly revealed.

Note first that the root vectors b_i, $i = 1, \ldots, n$, are linearly independent. Note also that the glue vectors $g_i = b_{i+n}$, $i = 1, \ldots, \ell$, are in the R-span of the root vectors. In fact, for each glue vector g_i there exists a smallest positive integer k_i (called the *characteristic* of g_i) such that

$$k_i g_i = -\sum_{j=1}^{n} c_{ij} b_j,$$

where each $c_{ij} \in Z$. If the g_i are Λ'-independent, these ℓ relations generate all others, in the sense that $\sum_{i=1}^{\ell} n_i g_i \in \Lambda'$ iff each n_i is a multiple of k_i — in fact, for $n_i \in Z$,

$\sum_{i=1}^{n+\ell} n_i b_i = 0$ would imply that $\exists m_1, \ldots, m_\ell \in Z$ such that

$$n_i = \begin{cases} \sum_{j=1}^{\ell} m_j c_{ji} & \text{for } i = 1, \ldots, n \\ m_{i-n} k_{i-n} & \text{for } i > n \end{cases} .$$

Now suppose for example that we remove the nodes corresponding to the roots $b_1, \ldots, b_\ell$. Then b_1 is in $Z\{b_i \mid i > \ell\}$ iff there are integers n_i such that $b_1 = \sum_{i=\ell+1}^{n+\ell} n_i b_i$. But then Λ' independence implies that there exist integers $m_{11}, \ldots, m_{1\ell}$ such that $1 = -\sum_{j=1}^{\ell} m_{1j} c_{j1}$, but $0 = -\sum_{j=1}^{\ell} m_{1j} c_{ji}$ for $i = 2, \ldots, \ell$.

Similar remarks hold when $b_2, \ldots, b_\ell \in Z\{b_i \mid i > \ell\}$. Letting M and C denote the Z-matrices with elements m_{ij} and c_{ij}, we get simply that $-MC = I$. Thus we have shown that if $\{b_{\ell+1}, \ldots, b_{\ell+n}\}$ is a basis for the gluing $\Lambda = \langle \Lambda', g_1, \ldots, g_\ell \rangle$, then the matrix C is invertible over the integers, and hence $|C| = \pm 1$.

The converse is also true, and its proof doesn't even require Λ'-independence. We have thus established a central result of this section:

Theorem 3.2.1 *Suppose the glue vectors are Λ'-independent. Let C denote the $\ell \times n$ matrix whose entries are c_{ij}. Then the gluing $\Lambda = \langle \Lambda', g_1, \ldots, g_\ell \rangle$ is redundant iff $\exists J \subset I_n$ such that $|C_J| = \pm 1$. In particular, $\{b_i \mid i \notin J\}$ is a basis of Λ iff $|C_J| = \pm 1$.*

(Of course, here the cardinality $\|J\|$ of J is ℓ, and C_J is the $\ell \times \ell$ submatrix of C obtained by retaining only those columns in C labelled by an index in J.)

Corollary 3.2.1 *Suppose $\exists J \subset I_n$ such that $|C_J| = \pm 1$. Then $\{b_i \mid i \notin J\}$ is a basis for Λ. (This is true even if the glue vectors fail to be Λ'-independent.)*

Note that it isn't required here that Λ and Λ' be integral or self-dual. Λ' doesn't have to be the direct sum of root lattices, and the g_i don't have to be weight vectors, but these are the cases of most interest.

Theorem 1 forms the foundation for the first method we shall discuss; as we shall see it is a very practical tool, especially for small ℓ. The $\ell \times n$ matrix C can be calculated effortlessly from Table 5. Table 5 should be used in conjunction with Tables 2 and 3.

Example 3.2.1 *D_n^+ for n even (see Example 6 for n odd).*

According to Table 5, $C = (\frac{n}{2}, \frac{n}{2} - 1, n-2, n-3, \ldots, 1)$ (since $k = \ell = 1$ C can be copied directly from the table). C has a component (i.e. a 1×1 determinant) equal to ± 1 — namely that one corresponding to r_n. Therefore, r_n can be expressed in terms of $r_i, i = 1, \ldots, n-1$, and [1], so D_n^+ is redundant, and its nth node should be removed (see Figure 8(a) for its true Coxeter diagram. Incidently we shall use that diagram to show in §4 that D_n^+ is self-dual for all $n \equiv 0$ (mod 4) — for $n \not\equiv 0$ (mod 4) D_n^+ fails to be integral).

Similar reasoning shows $A_n^* = A_n[1]$ and $A_{24}[5]$ are both redundant (in both cases the first node is removed).

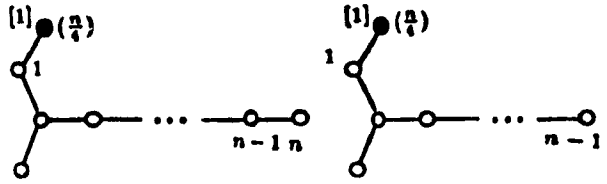
Example 3.2.2 *Show $D_4[1] = D_4[2] = D_4[3] = Z^4$ (where equality here is congruence).*

It is trivial to verify that these 3 gluings are all redundant, with an identical "true" Coxeter diagram (see Figure 8(b)). Thus they all have the same Gram matrix and thus are all congruent (see Theorem 2.3.3). We know

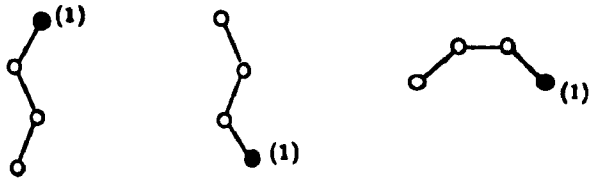
Table 5: The Basis Method

	Glue Vectors	Characteristic k_i	$[C_{i1}, \ldots, C_{in}]$
A_n	[0]	1	[0,...,0]
	$[i]$	$\frac{n+1}{(i,n+1)}$	$\frac{1}{(i,n+1)}[i, 2i, \ldots, ji, j(i-1), \ldots, j]$
	$i = 1, \ldots, n$		$(j = n + 1 - i)$
D_n	[0]	1	[0,...,0]
	[1]	2 if n even	$[\frac{n}{2}, \frac{n}{2} - 1, n-2, n-3, \ldots, 1]$
		4 if n odd	$[n, n-2, 2n-4, 2n-6, \ldots, 2]$
	[2]	2	[1,1,2,2,...,2]
	[3]	2 if n even	$[\frac{n}{2} - 1, \frac{n}{2}, n-2, n-3, \ldots, 1]$
		4 if n odd	$[n-2, n, 2n-4, 2n-6, \ldots, 2]$
E_7	[0]	1	[0,...,0]
	[1]	2	[3,6,9,12,8,4,7]
E_6	[0]	1	[0,...,0]
	[1]	3	[4,2,0,1,2,3]
	[2]	3	[2,4,6,5,4,3]
Z^n, E_8	[0]	1	[0,...,0]

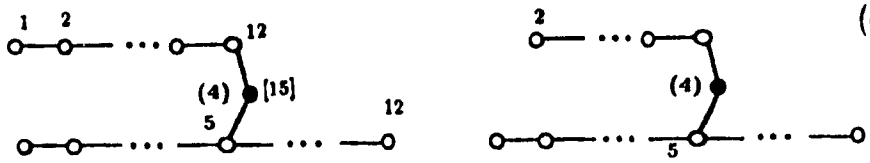
Figure 8: The Coxeter Diagrams of the Examples of the First Method



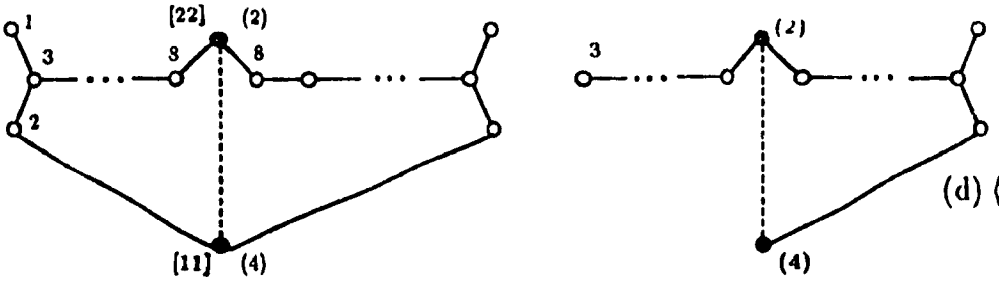
(a) D_n^+ (improper and proper Coxeter diagrams)



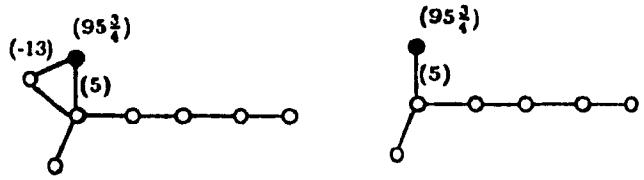
(b) The proper diagrams for $D_4[i]$



(c) $\langle A_{12}, A_{12}, [15] \rangle$ (improper and proper diagrams)



(d) $\langle D_8, D_8, [11], [22] \rangle$ revisited



(e) An alternate, superior choice of glue generator for D_n^+, n odd

$D_4^+ = D_4[1]$ is self-dual. From Table 1 we see there is only 1 self-dual integral lattice of dimension 4. Since Z^4 is also Type I, it must also be congruent to those 3 gluings.

Witt's Theorem tells us that if each independent glue vector g_i has norm 1 or 2 (as was the case above), then the gluing of root lattices can be replaced with (i.e. is congruent to) a mere direct sum of (different) root lattices. This rarely happens, except for small k. The converse must be slightly weakened: If a gluing of root lattices is congruent to a direct sum of root lattices, then it is possible to choose a set of glue vectors g'_j which span all glue vectors (i.e. the g'_j also are generators of the glue group), and all g'_j have norm 1 and 2. These g'_j may or may not equal the original set g_i of generators — i.e. not all g_i may have norm 1 or 2.

Example 3.2.3 $\langle A_{12}, A_{12}, [15]\rangle$, *one of the Niemeier lattices (a preponderance of these examples will involve Niemeier lattices — i.e. 24-dimensional Type II lattices. This is partly because these represent the greatest single triumph of the gluing process, but mostly because their glue vectors are explicitly known (see Table 4)).*

Here, the glue vector [15] has characteristic equal to $lcm(12+1, 12+1) = 13$. *From the chart 13[15] yields the matrix*

$$C = (1, 2, \ldots, 12, 5, 10, \ldots, 40, 32, \ldots, 8),$$

which has a ± 1 *in the first component. Thus* $\langle A_{12}, A_{12}, [15]\rangle$ *is also redundant.*

Similarly, $\langle A_{11}, D_7, E_6, [111]\rangle$ is redundant ([111] has characteristic lcm$(11+1, 4, 3) = 12$), as are $\langle A_{17}, E_7, [31]\rangle$ and $\langle A_{15}, D_9, [21]\rangle$ (the first node is to be removed in all of these).

Example 3.2.4 $\langle D_8, D_8, [11], [22]\rangle$

$$\textit{Here, } C = \left(\begin{array}{cccccccc|cccccccc} 4 & 3 & 6 & 5 & 4 & 3 & 2 & 1 & 4 & 3 & 6 & 5 & 4 & 3 & 2 & 1 \\ 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 \end{array}\right).$$

$$\textit{Note that } |C_{\{1,2\}}| = \begin{vmatrix} 4 & 3 \\ 1 & 1 \end{vmatrix} = 1 \textit{ and } |C_{\{1,3\}}| = \begin{vmatrix} 4 & 6 \\ 1 & 2 \end{vmatrix} = 2.$$

Thus we can simultaneously remove nodes 1 and 2 (i.e. root vectors $b_1 = "r_1 \oplus \vec{0}"$ and $b_2 = "r_2 \oplus \vec{0}"$), but not nodes 1 and 3. It can be shown that a total of 8 pairs of nodes may be removed: namely, 1 and 2; 2 and 4; 9 and 10 (i.e. $b_9 = "\vec{0} \oplus r_1"$ and $b_{10} = "\vec{0} \oplus r_2"$); 10 and 12; 1 and 10; 9 and 2; 2 and 12; and 10 and 4. Figure 8(c) illustrates the removal of 1 and 2; Figure 7(c) illustrates 2 and 12.

In all of the examples given thus far, the removal of nodes did not disconnect the originally connected Coxeter diagram. This pattern will continue. For one thing, a disconnected diagram implies the gluing is decomposable into a direct sum of other gluings. Secondly, it is possible to prove, at least when Λ' is a sum of root lattices and Λ is integral, that each component of that decomposition includes unit vectors (though these unit vectors may not be represented by a node in the diagram). Thus, not only would Λ be decomposable, but its components would include Z. This situation — a connected diagram being disconnected by the removal of redundant nodes — is not only rare, but also uninteresting, as there would clearly be more convenient and revealing ways to express Λ than as the given gluing.

In none of the examples given has the demand that the glue vectors g_i be Λ'-independent presented any problems. However the demand is not always as trivially satisfied. Λ'-independence of the generators g_i implies that the total number of glue vectors (i.e. the order of the glue group) of the gluing must equal the product of the characteristics of each generating glue vector g_i. (The total number of glue vectors can be computed by Theorem 2.7.1. For Λ self-dual, it is $\prod \Lambda_i$ — otherwise this is just an upper bound.) Consulting Table 4 we quickly find that although most of the given choices of g_i are Λ'-independent, not all are. For example, the glue vectors [240], [501] and [053] of $A_9^2 D_6$ have characteristics 5, 2 and 2 respectively, and apparently generate a glue group with 20 ($=\sqrt{10^2 4}$) elements (see Table 4), and so are $(A_9^2 D_6)$-independent. Similarly the given glue vectors for the 12 Niemeier lattices above it in Table 4 all are Λ'-independent. On the other hand, $9^3 \neq 27$, so choosing as the generators for the A_8^3 Niemeier lattice the glue vectors [114], [411] and [141] is not acceptable for our purposes (see Example 1, §3).

Any finite abelian group (e.g. the glue group) can be expressed as the direct product of cyclic groups C_i. This implies that, given any gluing of Λ', it is possible to choose Λ'-independent generators of the glue vectors. Well-documented procedures (involving "column and row operations" over

the integers) exist which can be used to derive Λ'-independent generators given any set of generators g_i (see Example 1 of the following section for a non-trivial example of this; the general theory is described on pp. 343-345 of HUN).

For small ℓ, it is also possible to perturb Theorem 1 to accommodate Λ'-dependent g_i, as the following example indicates.

Example 3.2.5 $\langle A_7^2 D_5^2, [1112], [1721] \rangle$ $g_1 = [1112]$ *and* g_2=*[1721] are not* $(A_7^2 D_5^2)$*-independent, as both* $8^2 \neq 32$ *and* $4g_1 + 4g_2 \in A_7^2 D_5^2$ *indicate. Here we have that* $C =$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7|1 & 2 & 3 & 4 & 5 & 6 & 7|10 & 6 & 12 & 8 & 4|4 & 4 & 8 & 8 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7|7 & 6 & 5 & 4 & 3 & 2 & 1|4 & 4 & 8 & 8 & 8|10 & 6 & 12 & 8 & 4 \end{pmatrix}$$

Instead of having to look for a C_J^J *invertible over* Z*, it suffices to look for one invertible over* $\frac{1}{2}Z$ *— i.e.* $|C_J^J|$ *could be* ± 2 *and still be acceptable. And while* $|C_J^J|$ *is never* ± 1*, there are some* J *for which* $|C_J^J| = 2$ *— e.g.* $J = \{1, 10\}$*. It is trivial to verify that*

$$M = \begin{pmatrix} -\frac{5}{2} & \frac{3}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

induces integers n_i, n_i' *satisfying* $b_1 = \sum_{i \neq 1,10}^{26} n_i b_i$*,* $b_{10} = \sum_{i \neq 1,10}^{26} n_i' b_i$ *(using the notation of Theorem 1), so we get that* $\langle A_7^2 D_5^2, [1112], [1721] \rangle$ *is redundant, and nodes 1 and 10 can be removed.*

Unfortunately, not all gluings are redundant. But even then a basis can sometimes be found.

Example 3.2.6 D_n^+ *for* n *odd (in which case* $D_n^+ = D_n^*$*).*

Here $C = (n, n-2, 2n-4, 2n-6, \ldots, 2)$*. As none of these entries are* ± 1*,* D_n^+ *is not redundant, at least as it currently stands. In other words, no node lies in the* Z*-span of the remaining nodes.*

Note however that 4 (the characteristic of [1]) is relatively prime to n *(the first entry in* C*). Thus, there exist integers* u *and* v *such that* $4u + nv = 1$*. Now if we replace* $g_1 = [1]$ *with* $g_1' = v[1] + ub_1$*, we see that* C *becomes*

$$C' = (vn + 4u, v(n-2), v(2n-4), \ldots, v2).$$

Now the first node may be successfully removed. Of course, g_1' is precisely as adequate as g_1 for generating the glue group. The only catch is that choosing new glue vector generators complicates the Coxeter diagram.

The case for D_7^+ is given in Figure 8(e) (there, $g_1' = 3[1] - 5b_1$). (Incidently, $|D_n^+| = \frac{1}{4}$ for n odd).

Similarly $A_{24}[5] = A_{24}[15]$, but $A_{24}[5]$ is redundant and $A_{24}[15]$ isn't. Several other examples along these lines can be found.

To generalize these examples, when the gluing isn't redundant it often is possible to make it redundant by choosing different g_i. Of course the new g_i should be Λ'-independent, and must generate the glue group.

Unfortunately, some gluings fail on a more fundamental level to be redundant. Examples include $A_1^2A_2^3[11111]$ and $E_7^2E_6^3[11111]$. These aren't redundant, for the characteristic of [11111] (namely, $\text{lcm}(2,2,3,3,3) = 6$) is larger than the characteristic of each [1] (which is 2 or 3). Thus, each entry in C is a multiple either of 3 or 2, so not only are none of them ± 1, but none are relatively prime to 6. Hence no alternate choice of g_1 will suffice.

Although these examples are integral, they fail to be self-dual (they fail the test given in the last section (see also Example 3 in the next section); a quick calculation shows their determinant is in fact 3). Every self-dual gluing I have considered either was redundant, or, by transforming the generators g_i as in Example 6, could become redundant. However, I have been unable to prove that this will always be the case.

At the beginning of the following section is given a variation of this method which is significantly more efficient (at least for the problem of finding the determinant, as opposed to a basis, of a gluing.

3.3 Alternatives: The Methods of Characteristics, and of GCD

The previous method involves calculating at most $\binom{n}{\ell}$ $\ell \times \ell$ determinants. This is trivial for small ℓ, as we have seen. If the search through these is successful, one finds a basis for the gluing Λ, from which, for example, $|\Lambda|$ can be quickly calculated using the techniques described in the following section.

One problem is that ℓ may be large (e.g. for one of the Niemeier lattices, $\ell = 23 = n - 1$, though for most Niemeier lattices $\ell \leq 3$), and since these $\ell \times \ell$ matrices aren't sparse, unlike those for $\Lambda_{\{i_1,\ldots,i_\ell\}}$, no tricks allow for the

quick calculation of their determinant. Also, there is a possibility that the searches described above (including those in Example 6) will fail and that a basis cannot be conveniently found. And lastly, there is the complication that the glue generators must be Λ'-independent.

Reminding ourselves that what is actually desired is merely $|\Lambda|$, the question naturally arises that perhaps the above method may try to do too much. Perhaps slightly moderating our ambitions will yield more positive results.

The two methods given in this section compute $|\Lambda|$ directly. The first is a significant improvement over the method of §2, although to find a basis the latter method is still required.

Theorem 2.7.1 implies that the total number $|G|$ of glue vectors equals $\sqrt{\frac{|\Lambda_1|\cdots|\Lambda_k|}{|\Lambda|}}$. Thus if we have some independent way of finding $|G|$, we can immediately deduce $|\Lambda|$

Theorem 3.3.1 *Suppose $g_1, \ldots, g_\ell$ are Λ'-independent and generate the glue group. Then Λ' is self-dual iff it is integral, and $k_1^2 \cdots k_\ell^2 = |\Lambda_1| \cdots |\Lambda_k|$, where k_i is the characteristic of g_i.*

Since k_i can be trivially computed using Table 5, and since $|\Lambda_i|$ can be read off from Table 2, Theorem 1 instantaneously enables us to determine if Λ is self-dual (if it isn't, we can use this argument to determine what $|\Lambda|$ is). All that is required is the Λ'-independence of the glue generators, which is also required by the method of §2.

Λ'-independence is trivially satisfied whenever $\ell = 1$. For example, we can trivially show that the only self-dual gluings of the form $A_n[k]$ are precisely $A_{k^2-1}[k]$ for $k = 2, 3, \ldots$ (for example $A_{24}[5]$). Also, the only self-dual lattice of the form $E_7^n[k_1 \ldots k_n]$ is $E_7^2[11]$ (which, incidently, is the only 14-dimensional indecomposable self-dual lattice), and there are no self-dual lattices of the form $E_6^n[k_1 \ldots k_n]$. For $\ell > 1$, to get Λ'-independence will in general require some straightforward manipulations (see pp. 343-345, HUN).

Example 3.3.1 *The Niemeier lattice $\Lambda = \langle A_8^3, [(114)] \rangle$*

Unfortunately $g_1 = [114]$, $g_2 = [411]$ and $g_3 = [141]$ aren't (A_8^3)-independent — not by a longshot. They all have characteristic 9, and a quick check reveals $3g_1 + 3g_2 + 3g_3 = 0$ (to see this and the other calculations here, use the glue group column in Table 3, so addition of glue vectors simply becomes

addition modulo 9). Now applying elementary row and column operations we get

$$\begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \\ 3 & 3 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ -9 & -9 & 9 \\ 0 & 0 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

The first operation operated on the columns, and so corresponded to a change of glue vector generators. In fact $g_3 \longrightarrow g_3' = g_1 + g_2 + g_3 = [666]$ *with a new characteristic of 3, while* g_1 *and* g_2 *were unaffected. Note also that* $g_1 - g_2 = [603]$, *so* $3g_1 - 3g_2 = 0$.

$$\begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \\ 3 & -3 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 9 & 9 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \\ 3 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \\ 3 & 0 & 0 \end{pmatrix}.$$

Again the first operation is the only column one, corresponding to $g_1' = g_1 - g_2 = [603]$ *for a new characteristic of 3, while* g_2 *and* g_3' *are left unchanged. Finally note that* $3g_2 = [333]$, *so* $6g_2 - g_3 = 0$. *We can proceed as above, or we may simply drop* g_3'. *Either way, we get that* g_1, g_2, *and* g_3 *can be replaced with* $g_1' = [603]$ *and* $g_2 = [411]$.

There are several ways we can verify their (A_8^3)*-independence. For example, looking at the second "component" of* $m[603] + n[411] = 0$ *gives us that 9 divides* n.

Of course, Theorem 1 now immediately tells us that this gluing is self-dual.

Theorem 3.3.2 *Let* $g_1, \ldots, g_\ell$ *be generators, not necessarily* Λ'*-independent. Then* Λ *fails to be self-dual if* $k_1^2 \cdots k_\ell^2 < |\Lambda_1| \cdots |\Lambda_k|$.

The final method I will discuss also computes $|\Lambda|$ directly. It always works (unlike the first method when the gluing isn't redundant), it doesn't require Λ'-independence in any form, but at least for smaller ℓ requires more effort than the previous two methods. We need first a preliminary result.

Let A be an $n' \times n'$ matrix, and let J and J' be subsets of $I_{n'}$ with elements $J_1 < J_2 < \cdots < J_n$ and $J_1' < \cdots < J_n'$, respectively (so $\|J\| = \|J'\| = n$). Define $A_{J'}^J$ to be the $n \times n$ submatrix of A satisfying $(A_{J'}^J)_{i,j} = A_{J_i J_j'}$ (so, for example, $|\Lambda_J| = |A_J^J|$).

Theorem 3.3.3 *Suppose $A = BC$, where B and C are $n' \times n''$ and $n'' \times n'$ matrices, respectively. Then*

$$|A^J_{J'}| = \sum |B^J_{J''}| \cdot |C^{J''}_J|,$$

where the sum is over all those sets $J'' \subseteq I_{n''}$ with n elements.

(Note that this formula closely resembles ordinary matrix multiplication — in fact it reduces to it when $n = 1$. Also, for $n' = n = n''$ this becomes the usual $|BC| = |B| \cdot |C|$. In the event that $n'' < n$, no such sets J'' can be found, and the sum, being over the empty collection, must vanish. This corresponds to the well-known fact that the determinant of a matrix not of full rank must be zero (A, and hence $A^J_{J'}$, can have rank at most n'').)

proof: Note that the ith column of BC is $\sum_j c_{ji}\vec{b}_j$. The determinant can be interpreted as a bilinear alternating form acting on column vectors. This implies

$$\det(\sum_{j_1}^{n''} c_{j_1 J'_1}\vec{b}^J_{j_1}, \sum_{j_2}^{n''} c_{j_2 J'_2}\vec{b}^J_{j_2}, \ldots) = \sum_{j_1}^{n''} c_{j_1 J'_1}\det(\vec{b}^J_{j_1}, \sum_{j_2}^{n''} c_{j_2 J'_2}\vec{b}^J_{j_2}, \ldots) = \cdots =$$

$$\sum_{j_1}^{n''} \cdots \sum_{j_n}^{n''} c_{j_1 J'_1} \cdots c_{j_n J'_n} \det(\vec{b}^J_{j_1}, \ldots, \vec{b}^J_{j_n})$$

$$= \sum_{j_1}^{n''} \cdots \sum_{j_n}^{n''} \epsilon(j_1, \ldots, j_n) c_{j_1 J'_1} \cdots c_{j_n J'_n} |B^J_{J''}|$$

where $J'' = \{j_1, \ldots, j_n\}$ and $\epsilon(j_1, \ldots, j_n)$ is the Levi-Civita symbol, equal to ± 1 or 0. Of course this last term just simplifies to the desired expression,

$$\sum_{J''} |C^{J''}_J| \cdot |B^J_{J''}|.$$

Now let's return to the gluing $\Lambda = \langle \Lambda_1, \ldots, \Lambda_k, g_1, \ldots, g_\ell \rangle$ (refer to the beginning of this chapter for the notation).

Let M be the $(n + \ell) \times m$ (non-integer) matrix

$$\begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n+\ell,1} & \cdots & b_{n+\ell,m} \end{pmatrix}.$$

Choose a basis for Λ and let M' be the corresponding generator matrix. This basis spans Λ, so there exists an integer matrix V such that $VM' = M$. Similarly, the root and glue vectors also span Λ, so there exists an integer matrix U such that $UM = M'$.

Let A be the $(n+\ell) \times (n+\ell)$ matrix MM^T, and let A' the $n \times n$ Gram matrix $M'M'^T$. Then $UAU^T = A'$ and $VA'V^T = A$.

The above theorem now implies $|\Lambda| = |A'| = \sum \iota_{JJ'} |A^J_{J'}|$, where the sum is over all $J', J \subset I_{n+\ell}$ with $\|J'\| = \|J\| = n$, and where $\iota_{JJ'} = |U_J| \cdot |(U^T)^{J'}| \in Z$. Therefore the greatest common divisor of all $|A^J_{J'}|$ must divide $|\Lambda|$.

Similarly, $|A^J_{J'}| = \iota'_{JJ'} = |V^J| \cdot |(V^T)_{J'}| \in Z$. Therefore $|\Lambda|$ must divide each $|A^J_{J'}|$, and hence their greatest common divisor.

Thus, we have proven:

Theorem 3.3.4 $|\Lambda| = gcd|A^J_{J'}|$, *provided* Λ *(i.e. A) is integral.*

In theory this requires $\binom{n+\ell}{\ell}^2$ $n \times n$ determinants to be taken, and then their gcd to be calculated. At least for self-dual Λ, in actual practice it seems only $\gcd|A^J_J|$ is required. Though it is geometrically obvious that $|\Lambda|$ must divide $\gcd|A^J_J|$, the reverse (to me, at least) seems far less clear. Armed merely with some general geometric considerations and a few specific examples, Lam proposed (prior to my discovery of Theorem 4) that perhaps $|\Lambda| = \gcd|A^J_J|$ always holds. Although I have been unable to prove (or disprove) this original formulation of the "gcd method", his intuition has since been vindicated, to a large extent, by Theorem 4.

Of course, by an earlier result each $|A^J_J|$ must be a perfect square if Λ is to be self-dual. Also, if a gcd of 1 is reached at any time during the calculation of the $\binom{n+\ell}{\ell}^2$ determinants, then (provided A is a Z-matrix) we may conclude that the gluing is self-dual. Since we're really only interested here in whether Λ is or isn't self-dual, for those reasons usually few determinants need be calculated (see, for example, the many examples given in LAM4, or the examples given below).

Theorem 4 is valid for any integral lattice, but for the cases we're interested in A is sparse and, hence, so are each $|A^J_{J'}|$. Thus these $n \times n$ determinants can be painlessly calculated (see the following section).

Example 3.3.2 D_n^+ *for even* n.

Choosing $J = J' = \{n\}$ yields $|A^J_{J'}| = 1$. Since D^+_n is integral, we need not go any further — the determinant of D^+_n must be a positive integer dividing 1, and hence must be 1.

Example 3.3.3 $E_7^2E_6^3[11111]$, *and* $A_1^2A_2^3[11111]$.

In both cases let $J = J' = \{n+1\}$ (i.e. remove the generator). Then $|A^J_{J'}|$ is simply 2^23^3, which is not a perfect square. The determinant of these two gluings must then both be a multiple of 3. Although we still don't know the exact value of the determinant at this early stage (in fact the determinant turns out to be exactly 3 in both cases), we do already know these gluings cannot be self-dual.

These two examples indicate the situations in which this gcd method can be prematurely terminated (in practice, it seems 2 or 3 determinants are necessary). They also hint that perhaps it *is* only necessary to consider $J = J'$.

Further (less trivial) applications of this method may be found in LAM4.

3.4 Calculating the Determinants

Two of the three methods considered in the previous two sections require the calculation of $n \times n$ determinants for large n. This section describes one way to do this.

Probably the quickest way in general to calculate the determinant of a large $N \times N$ matrix is by making it upper triangular by applying elementary row operations to the matrix. This amounts to making about $N^3/3$ divisions and multiplications, which can become unmanageable for large N.

The matrices considered here are fortunately of a rather special form: they're sparse (i.e. most of their entries are 0). Stunning computations are possible with these. Even in 1968 solving systems of linear equations of order 5000 was commonplace, and today this has increased a thousand-fold (see p. (v), PIS). All that is required is that the relevant matrices be sparse.

Our matrices are not only sparse, but they have a bandwidth of 3 almost everywhere. This has enabled Lam (see LAM4) to come up with an elegant graphical means of computing their determinants. Although in the computer age his technique is almost certainly unnecessary, aesthetic considerations demand its inclusion here.

The determinant of an $N \times N$ matrix A can be written as

$$|A| = \sum_{\pi \in S_N} (-1)^{o(\pi)} \prod_{i=1}^{N} A_{i\pi(i)},$$

where $o(\pi) = 0$ or 1 depending on whether π is, respectively, an even or odd permutation. Choose any index $j \in I_N$. Decompose π into a product of disjoint cycles, and let $C(\pi) = (j, j_1, \ldots, j_{\ell-1})$ denote the cycle containing j. Calling π^* the permutation $\pi / C(\pi)$, we get

$$|A| = \sum_{\pi \in S_N} \Gamma(C(\pi))(-1)^{o(\pi^*)} \prod_{i \in J} A_{i\pi^*(i)} = \sum_{C} \Gamma(C)|A_J^J| \text{ (call this (*))}$$

where $J = I_N \setminus \{j, j_1, \ldots, j_{\ell-1}\}$, $\Gamma(C) = A_{jj_1} A_{j_1 j_2} \cdots A_{j_{\ell-1} j}(-1)^{\ell+1}$, and where the final sum is over all cycles C containing j. A_j^J is the submatrix described in the previous section. This is called "expanding about node j" for reasons soon to be apparent.

For example, this formula can be used to get the recursions

$$|A_n| = 2|A_{n-1}| - (-1)^2|A_{n-2}| \ , \ |D_n| = 2|D_{n-1}| - (-1)^2|D_{n-2}|$$

and the equations

$$|E_6| = 2|D_5| - (-1)^2|A_4| \ , \ |E_7| = 2|D_6| - (-1)^2|A_5| \ , \ |E_8| = 2|D_7| - (-1)^2|A_6|$$

which can be solved to yield the values for the determinants given in Table 2. In the following two examples I explicitly show the calculations involved in finding these kinds of recursions or equations.

Example 3.4.1 D_{2n}^+ *(see Figure 8(a))*

We will expand about the glue vector [1] (this is a natural choice, as the resulting graphs will correspond to ordinary root lattices whose determinants we can read off from Table 2).

As always, the 1-cycle contributes a non-zero term: here

$$\Gamma(C) = A_{[1][1]}(-1)^{1+1} = 2n/4.$$

The graph that remains is precisely that of D_{2n-1}. We may now read off from Table 2 that $|D_{2n-1}| = 4$.

The only other cycle in the Coxeter diagram containing [1] is $C = ([1]1)$. *Then* $\Gamma(C) = (-1)(-1)(-1)^{2+1} = -1$. *The diagram that remains after removing the nodes in* C *is precisely that of* A_{2n-2} — *and Table 2 tells us* $|A_{2n-2}| = 2n-1$.

Thus $|D_{2n}^{+}| = 2n/4 \cdot 4 + (-1)(2n-1) = 1$.

Example 3.4.2 *We shall verify* $\langle D_8^2, [11], [22] \rangle$ *is self-dual by using the basis given in Figure 7(c), and by expanding about node* $j = [11]$. *The only cycles* C *with non-zero* $\Gamma(C)$ *are the 5 loops in the Coxeter diagram shown in Figure 7(d), together with the trivial cycle of length 1.*

First consider $C = (j)$. $\Gamma(C) = 4 \cdot (-1)^2 = 4$. *Throwing away* $\{j\}$ *(and the segments containing it) disconnects the diagram into two pieces. These pieces are indistinguishable from* A_{12} *(thanks to the good behaviour of [22]) and* A_3, *so* $|A_J^J|$ *becomes simply* $(12+1) \cdot (3+1) = 52$. *Thus this term contributed* $4 \cdot 52 = 208$.

There are 3 2-cycles in the diagram passing through [11]. One yields $\Gamma(C) = (-1)(-1)(-1)^3 = -1$ *and* $|A_J^J| = (11+1)(3+1) = 48$. *Another has* $\Gamma(C) = (-1)(-1)(-1)^3 = -1$ *and* $|A_J^J| = (12+1)(2+1) = 39$. *Finally, there is* $\Gamma(C) = (-1)(-1)(-1)^3 = -1$ *and* $|A_J^J| = (7+1)(4+1)(3+1) = 160$. *All other 2-cycles have* $\Gamma(C) = 0$ *and so can be ignored. These contribute* $-1 \cdot 48 - 1 \cdot 39 - 1 \cdot 160 = -247$.

The only other non-zero terms come from a 9-cycle.

$$\Gamma(C) = (-1)^8(+1)(-1)^{10}, |A_J^J| = (4+1)(3+1) = 20.$$

This cycle can be traversed both clockwise and counterclockwise (unlike for 2-cycles, these traversals are different — i.e. $C \neq C^{-1}$*), so we have an additional factor of 2.*

Therefore $|\Lambda| = |A| = 208 - 247 + 2 \cdot 20 = 1$ *and* Λ *is self-dual.*

Similarly, the determinants of all such 'redundant' gluings (see §2) can be computed.

An additional complication is introduced into the calculation of determinants $|A_{J'}^J|$ for $J \neq J'$. Of course, equation $(*)$ still holds; the problem is with the diagrams and their interpretation.

$J \neq J'$ simply implies $A_{J'}^J$ is no longer symmetric. This suggests we build our diagram out of directed line segments: e.g. represent the situation $(A_{J'}^J)_{ij} = -1$ and $(A_{J'}^J)_{ji} = 0$ with a segment linking nodes i and j, directed

from i to j. Note that in general a node label j may refer to a different node (i.e. basis vector) depending on whether you want a node at the tail of a segment (i.e. a *row* index of A, like J_j) or one at the head (i.e. a *column* index of A, like J'_j). All this talk of labels and row and column indices may seem out of place here --- after all, the method used in the above examples proceeded quite independently of any choice of node labels, etc. Indeed, this was one of its most appealing features. But equation (*), which lies at the heart of this method, presupposes such a choice. For the symmetric case $|A^J_J|$ we can (and did) completely ignore all this as irrelevant to the calculations, because any relabelling of the nodes induces identical permutations of the row and column indices, leaving $|A^J_J|$ unchanged. Equally important, the Coxeter diagram is also unaffected. But in the nonsymmetric case it is no longer desirable to permute the row and column indices identically, as we shall see. Because of this, $|A^J_{J'}|$ can be affected (though only trivially — it can change sign), but most relevant, the diagram and our notions of cycles and nodes are profoundly affected: the practicality of the method depends on the relabelling, for the labelling we inherit by default is hardly a satisfactory one in general.

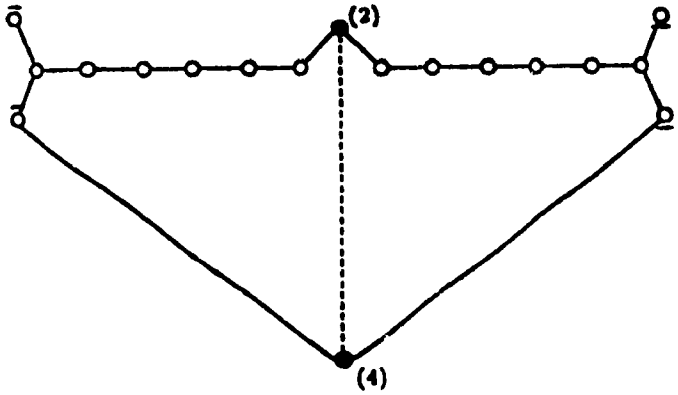
We shall rearrange the rows and columns of $A^J_{J'}$ so that, as much as possible (i.e. for all nodes in $J' \cap J$), the ith row and the ith column both refer to the same vector/node. This will at most change the sign of $|A^J_{J'}|$, so will have no influence on the final gcd calculation.

In particular, construct the diagrams as follows: represent by "o" (as usual) all indices in $J' \cap J$. Represent by "$\triangle$" each index in $J \setminus J'$, and by "$\triangledown$" those in $J' \setminus J$. Connect these with directed or undirected segments in the manner outlined above. For instance, all segments with an endpoint at a "$\triangle$" will be directed away from that node; all with one at a "$\triangledown$" will be directed into it.

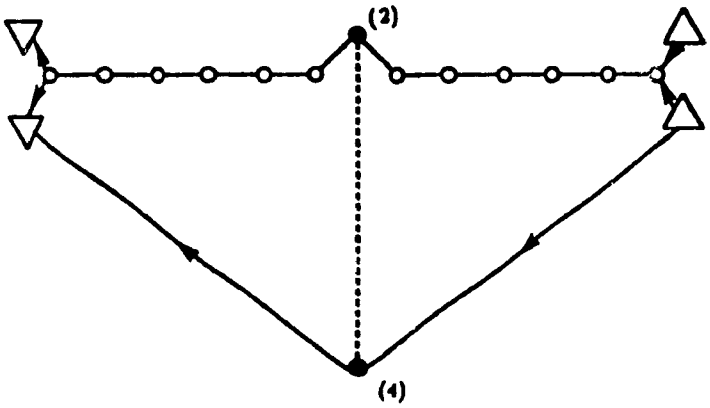
For example, consider again $\langle D^2_8, [11], [22] \rangle$. See Figure 9(a). Suppose J consists of all nodes except for the two shown with a bar above them, and suppose J' consists of all but those underlined. We would represent this as in Figure 9(b).

Note that in general there will be at most ℓ nodes looking like "$\triangle$", and an equal number like "$\triangledown$". Pair these off, and treat each pair as a single node, as was done in Figure 9(c). There are different ways of doing this - some ways are better than others. Figure 9(d) gives the single (and less desirable) alternative to Figure 9(c) (among other things it is non-planar —

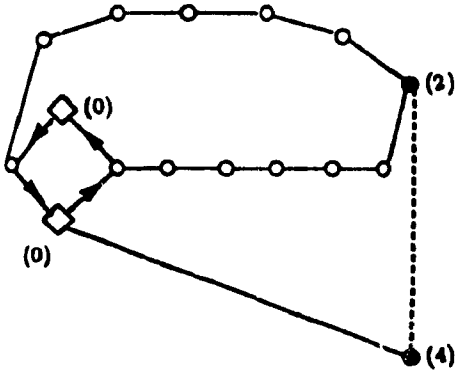
Figure 9: Coxeter Diagrams When $J \neq J'$



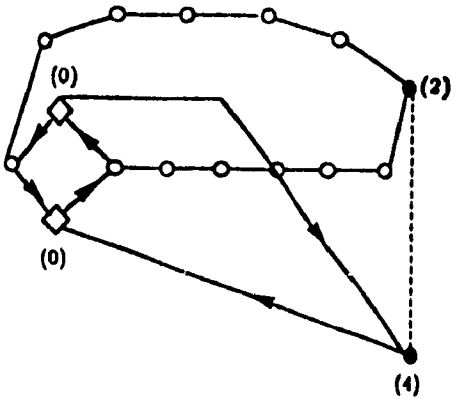
(a) Step 1: Locating the J and J' nodes



(b) Step 2: Introducing the triangles and arrows



(c) The simplest (completed) Coxeter diagram



(d) The nonplanar alternative

you must ignore the intersection of two of its lines).

This pairing refers to the choice of one index (say j) to refer to two different vectors. One of these vectors, the one corresponding to the "$\triangle$" in the pair, is referenced any time j is used as a row label; the other, the "$\triangledown$", corresponds to j as a column label. For the nodes like "o", these two vectors are equal; for the nodes like "$\diamond$" these vectors are different.

The use of equation $(*)$ now proceeds as before. The resulting Coxeter diagrams are more complicated than in the undirected ones, but the presence of directed segments reduces significantly the numbers of cycles.

Example 3.4.3 *Consider the diagram in Figure 9(c) of* $\langle D_8^2, [11], [22] \rangle$. *We shall expand about the upper "$\diamond$".*

Conveniently, there are only two cycles through this node. One is the cycle of length 1, of course, but here the corresponding term in (*) vanishes because the node has "norm" 0. The remaining cycle is of length 14. For it, $\Gamma = (-1)^{14}(-1)^{14+1} = -1$. *The resulting 2×2 determinant is* $\left|\begin{smallmatrix} 0 & -1 \\ -1 & 4 \end{smallmatrix}\right| = -1$. *Therefore* $|A^j_{j'}| = (-1)(-1) = +1$.

Most other examples are a little more complicated, and may require expanding about a second node to simplify the additional determinants that arise.

4 Compact Lattices

4.1 The Problem and Its Physical Context

In this chapter (at least in §§1-3) we aren't concerned directly with the lattices studied in the earlier chapters, but rather with their quotient — i.e. with compact or toroidal lattices. In Chapter 1, §5 we discussed a certain parametrization of the various string theories. However, any single string theory could be represented by different sets of parameters — we shall discuss this in the present chapter. Mathematically our task in this chapter will amount (as we shall see) to finding all bases of these toroidal lattices.

The boson fields X^B in a (closed) string theory are defined on the 2-dimensional manifold (σ, t). Since σ is the string parameter, σ and $\sigma + \pi$ represent the same point. X^B is itself not observable, so it isn't necessary for it to be well-defined in this sense (i.e. be periodic). Nevertheless its behaviour under the mapping $\sigma \longrightarrow \sigma + \pi$ is still of major importance, and the phase gained yields the "boundary conditions" of the field. As was mentioned in §5 of Chapter 1, the $(L+R)$-tuple $\tilde{w}$ characterizes these boundary conditions:

$$X^B(\sigma + \pi, t) = \exp(-2\pi i \epsilon^B \tilde{w}^B) X^B(\sigma, t)$$

where $\epsilon^A = \pm 1$ for left/right movers, and where $B = 1, 2, \ldots, L + R$. We found that we needed several boundary conditions (for the GSO projection); each choice of boundary condition corresponds to solutions living in a different Hilbert space, called a sector, and labelled by a q-tuple k.

$\tilde{w}(k) = (w_+^1, w_+^2, \ldots, w_+^L; w_-^1, \ldots, w_-^R)$ is the phase of the sector labelled with $k = (k_1, \ldots, k_q)$. It turns out to be nearly linear in k: $\tilde{w}(k) = \sum_{a=1}^{q} k_a \tilde{w}_a - \tilde{w}''(k)$, where (by convention) $0 \leq w_{a\pm}^A < 1$, and where $w_\pm''^A(k)$ is the integer chosen so that $w_\pm^A(k)$ is similarly in $[0, 1)$.

In addition, each $w_\pm^A(k)$ is rational, so we can write $w_\pm^A(k) = \frac{r_\pm^A(k)}{n_a}$, with n_a being the smallest common denominator of these components of $\tilde{w}_a$ (the consequences of selecting n_a to be a larger common denominator will soon be clear — essentially speaking, either simplicity or information is lost). Also, the integers k_a are in $[0, n_a)$. In short, each $\tilde{w}(k), \tilde{w}_a \in ([0, 1) \cap Q)^{L+R}$, and $k \in Z_{n_1} \times Z_{n_2} \times \cdots \times Z_{n_q} \equiv \prod_{a=1}^{q} Z_{n_a}$. The phases $\tilde{w}(k)$ give us the boundary conditions and are in this sense physically significant: changing them will change the solutions and hence will in general change the theory (except for

global rotations SO(L)× SO(R), since $\tilde{w}(k)$ enters into measurable quantities only in dot products, and except for translating the phases by integers, because of the factor $2\pi i$). In this chapter we'll consider redundancies in the characterization of string theories given, for example, in LAM3.

Different choices of parameters like $\tilde{w}_a$, R, L and q (usually) amount to specifying different string theories. However, only some of these have hopes of adequately representing nature (they must be modular invariant, etc.). For example, there must exist integers F_a (associated with the fermionic phase $F(k)$ of the vacuum $|\Omega\rangle_k$) satisfying:

(L) $2\tilde{r}_a \cdot \tilde{T} + n_a F_a \in 2Z$;

(Q) $\tilde{r}_a^2/n_a + n_a F_a \in 2Z$; and

(O) $\tilde{r}_a \cdot \tilde{r}_b / D_{ab} \in Z$, for all $a \neq b$.

Here, $\tilde{T} = (\frac{1}{2}, \ldots, \frac{1}{2})$, $D_{ab} = (n_a, n_b)$, and the metric determining the dot products is $G^{L,R}$. Denote (L), (Q) and (O) collectively by $\mathfrak{R}$ — they, together with (N) (see §5, Chapter 1), amount to ensuring the resulting theory be modular invariant.

By the phase "lattice" I mean the structure $\Lambda = \{\tilde{w}(k) \mid k \in \prod Z_{n_a}\}$. It's not a lattice as we defined the term in chapter 2 — rather, it's the quotient of the lattice $Z\langle\tilde{w}_a\rangle$ with its sublattice $Z^{L,R}$. Λ is thus a kind of discrete torus. The twist parameters $\tilde{w}_a$ uniquely determine Λ, and in fact constitute what we'll call a basis for it (provided the q vectors are linearly independent in a sense we'll specify in the next section).

Now, appropriate transformations of the $\tilde{w}_a$'s may merely "rephrase" a particular string theory. Changing the $\tilde{w}_a$'s doesn't necessarily change the $\tilde{w}(k)$'s, provided we also reshuffle our sector labels k. More precisely, suppose we have two different sets of twist parameters, $\{\tilde{w}_a | a = 1, \ldots, q\}$ and $\{\tilde{w}'_b | b = 1, \ldots, q'\}$, with denominators n_a and n'_b respectively. These both help define different theories, but under certain situations we may expect these theories to be only superficially different — isomorphic, in some sense. In particular, consider the possibility that there is a $q \times q'$ Z-matrix V such that $\tilde{w}'(V(k)) = \tilde{w}(k)$ $\forall k$. (As with the case of lattice congruence or similarity in §3, $q \neq q'$ is a possibility, corresponding to an embedding or projection.) V is the "reshuffling" mentioned above.

Whether V is itself one-to-one or onto is irrelevant. V induces a lattice transformation $\hat{V} : \Lambda \longrightarrow \Lambda'$ given by $\alpha \longrightarrow \tilde{w}'(V(\tilde{w}^{-1}(\alpha)))$. $\hat{V}$ must be well-defined (i.e. independent of the choice of $\tilde{w}^{-1}(\alpha)$ if $\tilde{w}$ isn't one-to-one), onto and one-to-one. In fact, from these properties of $\hat{V}$ we can conclude that

the set of points that comprise Λ equals that of Λ', and $\hat{V}$ can be interpreted as the identity on that set. Also, if the twist parameters $\tilde{w}_a$ are linearly independent, the condition on $\hat{V}$ reduces to the simpler condition that the restriction $V : \prod Z_{n_a} \longrightarrow \prod Z_{n_{b'}}$ be one-to-one and onto.

The reason we must have $\tilde{w}' \circ V = \tilde{w}$ (or, equivalently, that the points in Λ be precisely those in Λ') is because, unlike the twist parameters $\tilde{w}_a$ or the sector labels k, the phases $\tilde{w}$ have real physical significance. We can't change them in general without changing the fields they correspond to. Except for the trivial changes obtained by adding integers, we are considering in this chapter only transformations for which, symbolically, $\tilde{w}'(k') = \tilde{w}(k)$.

If such a V does exist, then we shall say that the linear map U sending $\{\tilde{w}_a\}$ to $\{\tilde{w}'_a\}$ is "reversible", and say that $\tilde{w}'_b$ is "attainable" from $\tilde{w}_a$, for lack of better terms — the differences in the two sets of twist parameters captured by this U can be undone by V. Reversibility is the generalization of invertibility appropriate here; it should succeed in revealing many of the redundancies inherent in the formulations of §5, Chapter 1.

The reversible U are precisely those mappings that can be interpreted as basis transformations of the phase lattice. The usual requirement of a basis transformation (e.g. one in a true lattice) is that it be invertible (for a lattice this amounts to $|U| = \pm 1$, we found). We can weaken this here to our notion of reversibility, thanks to the fact that Λ is the quotient of lattices. If U^{-1} exists then $V = U^{-1}$ works, but there is little justification for insisting upon the (seemingly natural) demand that V always equal U^{-1}.

We are being more general here than we may appear to be. For example, consider any function $F : \prod Z_{n_a} \longrightarrow \prod Z_{n'_b}$ satisfying $\tilde{w}' \circ F = \tilde{w}$. Then we can wlog take F to be linear, and represent it by a matrix V. (To see this, use the fact that $\tilde{w}$ and $\tilde{w}'$ are linear functions of k and k', mod 1. In particular, $\tilde{w}'(F(k_1+k_2)) = (\tilde{w}' \circ F)(k_1+k_2) = \tilde{w}(k_1+k_2) = \tilde{w}(k_1)+\tilde{w}(k_2) = (\tilde{w}' \circ F)(k_1) + (\tilde{w}' \circ F)(k_2) = \tilde{w}'(F(k_1)) + (\tilde{w}'(F(k_2)) = \tilde{w}'(F(k_1) + F(k_2))$.)

We will find out in the next section that attainability is an equivalence relation. There we will express reversibility in a much more appealing matrix notation. In particular, we will show that it will be sufficient, given some matrix $\underline{R}$, to find all Z-matrices U, V such that $\underline{R} = \underline{R}UV \pmod{n}$.

After solving this in §3, I will discuss how well these basis transformations respect the modular invariance relations $\mathfrak{R}$ given earlier.

Perhaps the most significant consequence of the following solution is the

development of convenient machinery allowing us to handle basis transformations in these compact or toroidal lattices. I have been unable to find this discussed in the mathematical literature.

In the final section of this chapter I will discuss briefly an unrelated problem: why the momentum lattice must be self-dual.

4.2 Restatement and Preliminary Results

Let n be any common multiple of $n_1, \ldots, n_q$, and set $m_a = \frac{n}{n_a}$. Define $\underline{R}$ to be the $(L+R) \times q$ Z-matrix given by

$$\underline{R} = n \begin{pmatrix} w^1_{1+} & w^1_{2+} & \cdots & w^1_{q+} \\ w^2_{1+} & \cdots & & \vdots \\ \vdots & & & \\ w^L_{1+} & & & w^L_{q+} \\ w^1_{1-} & & & \vdots \\ \vdots & & & \\ w^R_{1-} & \cdots & & w^R_{q-} \end{pmatrix} = \begin{pmatrix} m_1 r^1_{1+} & \cdots & m_q r^1_{q+} \\ \vdots & & \vdots \\ m_1 r^L_{1+} & & \\ m_1 r^1_{1-} & & \\ \vdots & & \vdots \\ m_1 r^R_{1-} & \cdots & m_q r^R_{q-} \end{pmatrix}$$

and let $\underline{R}_a$ denote its ath column. Define $N = \text{diag}(n_1, \ldots, n_q)$,

$$\tilde{R}(k) = n \begin{pmatrix} w^1_+ \\ \vdots \\ w^R_- \end{pmatrix} \text{ and finally } k = \begin{pmatrix} k_1 \\ \vdots \\ k_q \end{pmatrix}.$$

In this matrix notation, we have

$$\tilde{R}(k) = \underline{R} \cdot k + n\tilde{w}''(k).$$

The components of $\tilde{R}$, $\underline{R}$, k and $\tilde{w}''(k)$ are in Z_n, Z_n, Z_{n_a} and Z respectively. Note that incrementing some k_a by n_a changes each component of $\underline{R} \cdot k$ by a multiple of n, so extending $\tilde{w}''(k)$ in the obvious manner allows us to consider all $k \in (Z_n)^q$. Though this extension will prove to be mathematically desirable, it does present some complications, for now $m_1 m_2 \cdots m_q$ "labels" (i.e. different k's) represent the same sector, i.e. the same boundary conditions.

When translated into this new notation, the relations (L), (Q) and (O) become:

(L′): $2\underline{R}_a \cdot \tilde{T} + nF_a \in 2m_a Z$;

(Q′): $\underline{R}_a^2/n + nF_a \in 2m_a Z$; and

(O′): $\underline{R}_a \cdot \underline{R}_b \in n\bar{D}_{ab}Z (= m_a m_b D_{ab} Z)$, where $\bar{D}_{ab} = (m_a, m_b)$, for all $a \neq b$ in Z_q.

The vacuum parameters used here have the same value as those used earlier in $\mathfrak{R}$. (L′), (Q′) and (O′) are completely equivalent to (L), (Q) and (O) : given any initial set of parameters $\tilde{w}_a$, L, R, etc., if we define $\underline{R}_a$, etc. as above, then (L), (Q) and (O) will be satisfied if and only if (L′), (Q′) and (O′) are.

These relations follow from the original three in $\mathfrak{R}$ by repeatedly making use of $\underline{R}_a = m_a \tilde{r}_a$ and $n = n_a m_a$. Denote these three relations also by $\mathfrak{R}$.

It is certainly not required that n be the *least* common multiple of $n_1, n_2, \ldots$ — the form of the above relations is independent of which multiple we take. This is not so with n_a and the original equations. There, if we allowed n_a to be *any* common denominator of the $w^A_{a\pm}$'s, new factors, complicating both sets of expressions, would have to be introduced to accommodate the change, for failure to introduce these factors would result in the constraints $\mathfrak{R}$ being weaker than those given above. Thus, simplicity dictates that we choose the n_a's to be minimal. However, no such argument applies to our choice of n — in other words, we can fix it from the start, so the question of how it transforms is answered trivially.

Note that incrementing any element of $\underline{R}$ or k by n results only in the elements of $\underline{R}k$ being multiples of n. These multiples are readily absorbed by $\tilde{w}''$. Similarly, the right hand sides of $\mathfrak{R}$ absorb the new terms created on their left hand sides by this change in $\underline{R}$ or k. Thus, we may interpret $\tilde{R}(k) = \underline{R}k + n\tilde{w}''$ in two new ways: as an equation over all $k \in Z$, or as an equation in the integers modulo n (i.e. the elements of $\tilde{R}$, $\underline{R}$, k and $\tilde{w}''$ are all in Z_n). In this latter interpretation, $n\tilde{w}'' \equiv 0$, so we have effectively linearized our equation:

$$\tilde{R}(k) = \underline{R}k \pmod{n}.$$

In this way we can ignore the complicated behaviour of $\tilde{w}''$ under the transformations U and V — the role of $\tilde{w}''$ is effectively handled by using modular arithmetic.

We shall use both these interpretations in what follows, but for the most part we'll restrict ourselves to mod n.

To repeat, we are concerned with solving the equation (written symbolically)

$$\tilde{w}'(k') = \tilde{w}(k).$$

That is, we are interested in lattice transformations U (taking $\tilde{w}_a \to \tilde{w}'_b$), and relabellings V (taking $k \to k'$), whose net effect is to leave the phases $\tilde{w}$ of each sector unchanged. This amounts to choosing a new basis for the 'lattice'.

Using the notation introduced earlier, we thus demand that $\underline{R}k = \underline{R}UVk$ (mod n). Since this must be satisfied by all k, we immediately get

$$\underline{R} = \underline{R}UV \pmod{n}.$$

Unfortunately, this does not imply $UV = I$ — a reversible transformation *need not be invertible* over Z_n, and especially not over Z. In addition, the choice of lattice basis transformation (i.e. a reversible U) does not uniquely determine the transformation law of k (i.e. V), or vice versa. There exist $q \times q$ matrices $E \neq 0$ (e.g. of the form $E = NA$) satisfying $\underline{R}E = 0$. All that is required is that $UV = I + E$ for one of these E.

Theorem 3.5 on p. 351 of HUN tells us that:

Theorem 4.2.1 *U is invertible over a commutative ring R with unity $\Longleftrightarrow$ $|U|$ is a unit (i.e. an invertible element) in R.*

U invertible over R means that the R-matrix U has an inverse V which is also a R-matrix. This has two corollaries.

Corollary 4.2.1 *U is invertible over Z iff $|U| = \pm 1$.*

This implies, for example, that basis transformations of (true) lattices must have determinant ± 1.

Corollary 4.2.2 *U is invertible over Z_n iff $|U|$ is relatively prime to n.*

For example, the matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

is invertible over Z_5 — its inverse is

$$\begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}.$$

But U doesn't have to be invertible over Z (in which case $E = 0$) or Z_n (in which case $E \equiv 0 \pmod{n}$). For example, it is quite possible for a perfectly acceptable U to satisfy $|U| \equiv 0 \pmod{n}$, in violent opposition to our two corollaries. The determinant, such a powerful tool (as Theorem 1 indicates) in both linear algebra and lattice theory and elsewhere, is useless in the following analysis — it's too "high-level". We noticed some changes in going from linear algebra to lattice theory (e.g. not all generating sets had a subset which was a basis); there are additional casualties in going from lattice theory to compact lattice theory.

There are two possible sources of this complication: (i) the inevitable consequences of our choice of notation, namely the many-to-one labelling caused by our extension of the range of k, as well as the replacement of the n_a with a common multiple n; and (ii), a "removable" consequence corresponding not to the new notation but to a poor selection of parameters. The former source will always be present; the latter is present only if in the original formulation (where $k \in \prod Z_{n_a}$, etc.) different k's (i.e. different sectors) can correspond to identical phases $\tilde{w}(k)$ — alternatively, that the $\tilde{w}_a$'s are, in some weakened sense, linearly independent. We can avoid this (to an extent indicated below) by demanding that the $\tilde{w}_a$'s be linearly independent. However, this is a needlessly strong restriction. The specific form of linear independence we require, which we will call (n_a)-independence, is that whenever there exist numbers $c_a \in Z_{n_a}$ satisfying $\sum c_a \tilde{w}_a \in Z^q$, then $c_1 = \cdots = c_q = 0$. (This is very analogous to the definition of Λ-independence given in the previous chapter.) For example, $\tilde{w}_1 = (\frac{1}{2}, \cdots, \frac{1}{2})$ and $\tilde{w}_2 = (\frac{1}{3}, \cdots, \frac{1}{3})$ are independent in this sense by not in the usual sense. (ii) applies for instance when we do not insist that n_a be the *least* common denominator of the components of $\tilde{w}_a$.

Normally (e.g. in a vector space), when vectors are linearly dependent we can express one in terms of the others, and thus reduce by at least one our set of vectors. This is not necessarily the case here. What we can always do (as can be seen from Example 1 in §3 of the previous chapter) is linearly combine the vectors $\tilde{w}_a$ to form new vectors $\tilde{w}'_a$, whose denominators n'_a are in general smaller than n_a. Provided we discard those vectors (if any) with

$n'_a = 1$, the new set of $\tilde{w}'_a$ are (n'_a)-independent. But $n' = n$, and in all other respects the theory specified by these $\tilde{w}'_a$ is identical to that given by the original $\tilde{w}_a$ — subject to the appropriate reshuffling of the labels k of course. Such a basis transformation is in fact "reversible".

In linear algebra, an invertible transformation acting on a set of linearly independent vectors will map them onto a second set of linearly independent vectors — i.e. invertible transformations preserve linear independence. Unfortunately, (n_a)-independence is not preserved by reversible transformations, and this will cause some problems.

For example, consider $n = 6$,

$$\underline{R} = \begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix}, U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \text{ and } V = \begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix},$$

so V "reverses" U. Then

$$\underline{R}' = \begin{pmatrix} 3 & 5 \\ 3 & 1 \end{pmatrix}.$$

Note that $n_1 = n'_1 = 2$ while $n_2 = 3$, and $n'_2 = 6$. Since $n_1 n_2 < n'_1 n'_2$, we know that, though the columns of $\underline{R}$ are (n_a)-independent (as is readily seen), the columns of $\underline{R}'$ cannot be (n'_a)-independent (indeed,

$$1 \cdot \begin{pmatrix} 3 \\ 3 \end{pmatrix} + 3 \cdot \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 18 \\ 6 \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

but $1 \not\equiv 0 \pmod 6$.)

$\prod n_i$ represents the number of sectors (counting "multiplicities" — i.e. different labels that correspond to the same sector) if we only allow $k_a \in Z_{n_a}$. This number will change as we change bases. The minimum value it can take (which is realized whenever the basis is (n_a)-independent) equals the total number of physically different sectors, i.e. sectors corresponding to different boundary conditions. Allowing, however, $k_a \in Z_n$ yields n^q total sectors (counting multiplicities), with many repetitions.

Thus, we can avoid the unpleasantries of (ii) by demanding from the start that different sectors should correspond to different phases, or equivalently, that the $\tilde{w}_a$ be (n_a)-independent. (But strictly adhering to this demand would force us to throw away reversible transformations like that given above.) Whenever (ii) is satisfied, we can find an equivalent theory for which (ii) is not, so (ii) is indeed only an inconvenience caused by poor foresight

when choosing the parameters, and doesn't correspond to some "physical" effect.

So we can avoid (ii) by demanding (n_a)-independence of the $\tilde{w}_a$, but we must recognize that we cannot maintain this demand if we consider (as we will) compositions of basis transformations (unless we restrict our attention to basis transformations that preserve (n_a)-independence — i.e. for which $n_1 \cdots n_q = n'_1 \cdots n'_{q'}$. This suggestion is particularly interesting, as we'll see later).

(i) is another story entirely. You cannot get something for nothing, it seems, and the mathematical simplifications resulting from choosing a common denominator n for all the components of all the $\tilde{w}_a$'s — namely, that the additions of different $\tilde{r}_a$'s, relevant when considering basis transformations, now become trivial — are directly responsible. We cannot, as was the case with (ii), assume the problem away. Our intuitions that the transformations U and V be invertible, that given one the other is uniquely determined, are violated here, but only by this translated version of the problem (and by (ii)). Within the confines of the original, physical phase lattice these intuitions are indeed valid, but the corresponding mathematics is much more awkward and artificial. We have merely substituted one problem for another, but our claim is that the new problem is much more easily tractable than the old. In particular, standard mathematics may now be used to help us find all U and V, as we'll see in the next section.

Incidently, we have diligently distinguished between q and q' i.e. between the number of basis vectors $\tilde{w}_a$ and the number of basis vectors $\tilde{w}'_b$. Consider the following example: take $n = 6$,

$$\underline{R} = \begin{pmatrix} 3 & 2 \\ 3 & 2 \end{pmatrix} \text{ and } U = \begin{pmatrix} 1 \\ 5 \end{pmatrix}.$$

Here $q = 2$, of course, but

$$\underline{R}' = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

has $n'_1 = 6$ and $q' = 1$. U is reversible (e.g. take $V = (2, 3)$). In addition the columns of $\underline{R}$ and those of $\underline{R}'$ are (n_a)- and (n'_a)-independent, respectively.

Due to examples like that given above, it seems unjustified to demand $q = q'$. Thus it seems that even the concept of dimension doesn't extend naturally to these compact lattices.

Indeed, for each prime p let $N(p)$ be the number of $i \in I_q$ such that p divides n_i. Define $q_<$ to be the maximum (taken over all p) of $N(p)$, and define $q_>$ to be the sum of all $N(p)$. Then $q \in [q_<, q_>]$, and for any $q' \in [q_<, q_>]$, there exists a reversible transformation sending $q \to q'$. (This is related to the fact that $Z_m \times Z_n = Z_{mn}$ iff m and n are relatively prime.)

Finally, it should be mentioned that attainability, which simply means that the two functions $\tilde{w}$ and $\tilde{w}'$ can be connected by one of these sector relabellings V, is an equivalence relation.

Theorem 4.2.2 *Attainability is an equivalence relation.*

> proof: For reflexivity take $U = I = V$. For symmetry just interchange the roles of U and V. For transitivity let $U'' = UU'$ and $V'' = V'V$.

4.3 The Solution

Given a Z-matrix $\underline{R}$, we must find all Z-matrices U and V such that $\underline{R}UV = \underline{R}$.

A frequently exploited technique in linear algebra is to decompose a matrix into products of particularly simple matrices. The most common example of this involves elementary column and row matrices (abbreviated em). We can take these to be precisely those matrices of the following forms:

$$A_j = \begin{pmatrix} 0 & & & & 1 & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ 1 & & & & 0 & & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix}$$

i.e. the identity matrix with rows (or columns) 1 and j interchanged (for readability we haven't displayed the off-diagonal zeros). This corresponds to a U that rearranges the basis vectors $\tilde{w}_a$ so that $\tilde{w}'_1 = \tilde{w}_j$, $\tilde{w}'_j = \tilde{w}_1$, and $\tilde{w}'_i = \tilde{w}_i$ for $i \neq 1, j$.

$$B_d = \begin{pmatrix} d & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

corresponding to multiplying the basis vector $\tilde{w}_1$ by the scalar d, and leaving all other basis vectors $\tilde{w}_a$ unchanged.

$$C = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

corresponding to $\tilde{w}'_1 = \tilde{w}_1 + \tilde{w}_2$, and $\tilde{w}'_i = \tilde{w}_i$ for $i \neq 1$; as well as its transpose.

$$D_{q'q} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & & 1 \\ 0 & 0 & 0 & \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

if $q < q'$, or its transpose if $q > q'$ (so $D_{q'q}$ consists of zeroes everywhere except on its diagonal where it's all 1's).

A_j, B_d and C are all $q \times q$ Z-matrices. $D_{q'q}$ is $q' \times q$. Usually slightly more general matrices are chosen by textbooks, and never $D_{q'q}$, but these are clearly necessary and sufficient for our purposes (see Theorem 2). For example, note that $C^{n-1} = C^{-1} \pmod{n}$ (over the integers rather than Z_n we must include C^{-1} in our list).

Before we try to decompose U into products of these em's, let's try to determine which of these em's are themselves reversible.

Theorem 4.3.1 *(a) Each A_j is reversible;*

(b) B_d is reversible iff $\exists b_1, \ldots, b_q$ such that $b_\ell \in Z_{n_\ell}$, $\sum b_\ell \underline{R}_\ell \equiv 0 \pmod{n}$, and d divides $b_1 + 1 \bmod n_1$;

(c) C is always reversible; and

(d) $D_{q'q}$ is always reversible if $q < q'$; if $q > q'$ it's reversible iff for each $i > q'$ there exist numbers $b_{i\ell} \in Z$ satisfying $\underline{R}_i \equiv \sum_{\ell=1}^{q'} b_{i\ell} \underline{R}_\ell \pmod{n}$.

(For example, (b) holds whenever d is relatively prime to n_1. It can also happen, though, for other d if the $\tilde{w}_a$ are (n_a)-dependent.)

proof: (a) and (c) are obvious, since they are invertible (and not merely reversible). (d) is almost as easy: if $q < q'$ take $V = D_{qq'}$; if $q > q'$ there is a trivial relationship between the last $q - q'$ rows of V and the numbers $b_{i\ell}$. So consider (b).

"$\Longleftarrow$": Suppose d divides $b_1 + 1$ and $\sum b_\ell \underline{R}_\ell \equiv 0 \pmod{n}$. Let e be an element in Z_{n_1} satisfying $ed \equiv b_1 + 1 \pmod{n_1}$. Define V by

$$V = \begin{pmatrix} e & & & \\ b_2 & 1 & & \\ \vdots & & \ddots & \\ b_q & & & 1 \end{pmatrix}.$$

"$\Longrightarrow$": Let b'_ℓ be the components in the first column of the V corresponding to the given B_d. Define the numbers b_ℓ by $b_1 = db'_1 - 1 \pmod{n_1}$ and $b_i = b'_i$ for $i > 1$.

Theorem 4.3.2 *Any matrix U over Z_n can be expressed as a finite product of the em's A_j, B_d, C and $D_{q'q}$.*

proof: Suppose U is $q' \times q$. If $q > q'$ let $U' = D_{qq'}U$; if $q < q'$ let $U' = UD_{qq'}$. From p. 339 of HUN we see that any square matrix over a Euclidean domain (like Z, but unfortunately not like Z_n) can be expressed as a finite product of the em's A_j, B_d and C (for Z and many other domains we must add C^{-1}). Applying this to the Z-matrix U' we get the product $U' = E_1 \cdots E_k$. Then we get for U either $D_{q'q}E_1 \cdots E_k$ or $E_1 \cdots E_k D_{q'q}$. These expressions, holding as they do over Z, must also hold, of course, over Z_n.

Theorem 4.3.3 *If $U = E_1 \cdots E_k$ is reversible (relative to $\underline{R}$), then E_1 is reversible (relative to $\underline{R}$), E_2 is reversible (relative to $\underline{R}E_1$), .., and E_k is reversible (relative to $\underline{R}E_1 \cdots E_{k-1}$).*

proof: $\exists V$ such that $\underline{R}E_1 \cdots E_k V = \underline{R}$. Then $\underline{R} = (\underline{R})E_1(E_2 \cdots E_k V) = (\underline{R}E_1)E_2(E_3 \cdots E_k V) = \cdots = (\underline{R})(E_1 \cdots E_{k-1})E_k(V)$.

Corollary 4.3.1 *U is reversible with respect to $\underline{R}$ iff there exist em's $E_1, \ldots, E_k$ such that $U = E_1 \cdots E_k$ (mod n) and each E_i is reversible with respect to $\underline{R}E_1 \cdots E_{i-1}$) (these E_i are given by Theorem 1).*

Thus we have characterized all reversible matrices (with respect to a given $\underline{R}$), solving the main problem of this chapter.

An example of a reversible matrix with determinant $\equiv 0 \pmod{n}$ can now be easily constructed. Suppose we have $n_1 = n_2 = 3$, $n_3 = 4$ and $n = 12$. Let

$$U = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Then U is guaranteed to be reversible (it reverses itself), even though its determinant is 0 (mod n).

What remains is to check for modular invariance. In particular, if $\underline{R}$ satisfies $\mathfrak{R}$, then will $\underline{R}'$ if $\underline{R}'$ is attainable from $\underline{R}$? The answer is surprising.

Theorem 4.3.4 *A_j, B_d and $D_{q'q}$ all satisfy $\mathfrak{R}$ (at least when they're reversible). C does if $n_1 = n_1'$, but otherwise might not.*

> proof: Obvious for A_j, B_d and $D_{q'q}$. Consider (c), where $n_1 = n_1'$. Then n_2 divides n_1, m_1 divides m_2 (so $D_{ab} = m_1$) and $m_1 = m_1'$. Let $F_1' = F_1 + F_2$. Then it is straightforward to verify that $\mathfrak{R}$ is satisfied.

Consider $\tilde{R}_1 = (1,1,3,1,2)$, $\tilde{R}_2 = (3,1,1,1,2)$, and $n_1 = n_2 = n$, so $m_1 = m_2 = 1$. Then $\tilde{R}'_1 = (0,2,0,2,0)$, so $n_1' = 2 = m_1'$. $\underline{R}$ satisfies $\mathfrak{R}$, as can be immediately verified, but $\underline{R}'$ cannot satisfy the relation (Q').

This result is surprising because it seems to say that modular invariance, which in one sense guarantees that the choice of a basis (of a certain 2-dimensional torus — see §4 of Chapter 1) shouldn't matter, is itself, in this sense, basis dependent. As suggested by Theorem 4 and the example, it is conceivable that modular invariance (i.e. the relations $\mathfrak{R}$) is preserved by those U that preserve (n_a)-independence. This has neither been proven nor disproven at the time of writing. (Obviously one way it could be proved is if it could be shown that any U which preserves (n_a)-independence can be expressed as a product of em's that individually preserve (n_a)-independence — i.e. if an analogue to the corollary can be found).

4.4 Self-Duality of the Momentum Lattice

(This section is closer in spirit to the previous chapter, but is added here to more closely equalize the lengths of Chapters 3 and 4.)

In Chapter 3 we considered bosonizing all fermions, and looked at the set $\{p\}$ consisting of the momenta of these bosons. This set turns out to be a shifted lattice $\Lambda + t$ — where Λ is a true (as opposed to compact) lattice. Λ is indefinite with metric G^{N_L,N_R}, for $N_L = 24 - d$, and $N_R = 12 - d$. It can be shown to be integral, and, finally, to be self-dual. In this section I will briefly discuss the demonstration that Λ must be self-dual.

Λ self-dual is supposed to follow from the modular invariance of the partition function $Z(\tau, \bar{\tau})$. We have

$$Z(\tau, \bar{\tau}) = \eta(\tau)^{-N_L}\eta(\tau)^{*-N_R}L(\tau, \bar{\tau}) \text{ and}$$

$$L(\tau, \bar{\tau}) = \sum_{q\in\Lambda} \exp(\pi i\tau(q_L + t_L)^2 - \pi i\bar{\tau}(q_R + t_R)^2 + 2\pi iq \cdot s).$$

$\eta(\tau)$ is the Dedekind eta function, and τ is the modular parameter. $\bar{\tau}$ is its complex conjugate, but is treated as an independent variable. τ lies in the fundamental region F=$\{z \in C \mid \text{Im } z > 0, -\frac{1}{2} < \text{Re } z \leq \frac{1}{2}, |z| \geq 1\}$. s is the vector such that $2(p - t) \cdot s$ is the fermionic number.

Z must be invariant under the modular group, and since the modular group is generated by the transformations $\tau \to \tau + 1$ and $\tau \to -\frac{1}{\tau}$, it suffices to show that Z is invariant under these two. Self-duality is supposed to follow from the invariance of Z under the second one.

Define as in LAM3

$$G(u, v|\tau, \bar{\tau}) = \sum_{q\in\Lambda} \exp(\pi i\tau(q_L + v_L)^2 - \pi i\bar{\tau}(q_R + v_R)^2 + 2\pi iq \cdot u).$$

(For example, $G(s, t|\tau, \bar{\tau}) = L(\tau, \bar{\tau})$.) Under $\tau \to -\frac{1}{\tau}$ G becomes

$$G(u, v| -\frac{1}{\tau}, -\frac{1}{\bar{\tau}}) = \sum_{q\in\Lambda^*} \exp(\pi i\tau(q_L - u_L)^2 - \pi i\bar{\tau}(q_R - u_R)^2 + 2\pi iq \cdot v)$$

after using the Poisson summation formula.

We know how η transforms under $\tau \to -\frac{1}{\tau}$ ($\eta(-\frac{1}{\tau}) = (-i\tau)^{\frac{1}{2}}\eta(\tau)$), so we know that if Z is to have a hope of being modular invariant, $G(u, v| -\frac{1}{\tau}, -\frac{1}{\bar{\tau}})$ must be a certain scalar multiple of $G(u, v|\tau, \bar{\tau})$ (at least for $(u, v) = (s, t)$).

One obvious way this can happen is if $\Lambda^* = \Lambda$ (and s and t are related in the obvious way) — i.e. if Λ is self-dual. It is conjectured that this is the only way this can happen; hence the conclusion that Λ must be self-dual.

Clearly this argument has not succeeded in rigorously establishing that Λ must be self-dual, and to my knowledge no such proof has been found. Nevertheless, it is a very desirable property of Λ for, as we have shown in Chapter 2, self-duality is a very strong constraint. Its usefulness in greatly restricting possible models of superstrings can be quite graphically seen in both LAM1 and LAM4, for example (they do this explicitly for $d = 8$).

While I haven't found a rigorous proof of its self-duality, I have perhaps made that conclusion even more plausible. My argument is that in the special case where s and t are in Λ (in which case we can wlog take them to be the zero vector), self-duality is forced. (String theory, however, is more interested in s and t being in $\frac{1}{2}\Lambda^*$.)

Theorem 4.4.1 *$\eta(\tau)^{-N_L}\eta(\tau)^{*-N_R}G(0,0|\tau,\bar{\tau})$ is modular invariant for $N_L = 24 - d$, $N_R = 12 - d$ iff Λ is self-dual.*

proof: We must have $(-i\tau)^{-N_L/2}(i\bar{\tau})^{-N_R/2}G(0,0|-\frac{1}{\tau},-\frac{1}{\bar{\tau}}) = G(0,0|\tau,\bar{\tau})$ for all $\tau \in$ F. From the previous argument these become

$$\sum_{q\in\Lambda^*} \exp(\pi i\tau q_L^2 - \pi i\bar{\tau}q_R^2) \text{ and } \sum_{q\in\Lambda} \exp(\pi i\tau q_L^2 - \pi i\bar{\tau}q_R^2)$$

Take $\tau = iy$ for some $y \in R$, so $\bar{\tau} = -iy$. Then the expressions become

$$\sum_{q\in\Lambda^*} \exp(-\pi y(q_L^2 + q_R^2)) \text{ and } \sum_{q\in\Lambda} \exp(-\pi y(q_L^2 + q_R^2)).$$

Note that every term in both these series is positive. Also, Λ is integral, so $\Lambda \subseteq \Lambda^*$. If $\Lambda \neq \Lambda^*$, the left equation must be larger than the right one, so they cannot be equal. Thus modular invariance requires here that $\Lambda = \Lambda^*$.

This argument seems to collapse for $(u,v) = (s,t)$ because we no longer have the positivity of each term. The right equation will have an additional factor of $(-1)^F$ in each term (F is the fermionic number), and for the left equation the additional factor will be a more complicated complex number.

REFERENCES

BOR Borcherds, R. E. (1985), The Leech Lattice, *Proc. R. Soc. Lond.* **398A**, 365-376.

BOU Bourbaki, N. (1968), *Groupes et Algebres de Lie*, Ch. 4-9, (Hermann, Paris).

CAR Cartan, Elie (1966), *The Theory of Spinors*, (Dover).

CHA Chapline, George (1985), Unification of Gravity and Elementary Particle Interactions in 26 Dimensions?, *Phys. Lett.* **158B**, 393-396.

CS Conway, J. H. and Sloane, N. J. A. (1988), *Sphere Packings, Lattices and Groups*, (Springer-Verlag).

DB Davies, P. C. W. and Brown, J., editors (1988), *Superstrings: A Theory of Everything?*, (Cambridge University Press).

DYS Dyson, Freeman (1983), Unfashionable Pursuits, *Math. Intelligencer* **5, No. 3**, 47-54.

FRE Freund, Peter G. O. (1986), *Introduction to Supersymmetry*, (Cambridge University Press).

GG Ginsparg, Paul and Glashow, Sheldon (May, 1986), *Phys. Today* **39**, 7-9.

GHMR Gross, David J., Harvey, Jeffrey A., Martinec, Emil and Rohm, Ryan (1985), Heterotic String, *Phys. Rev. Lett.* **54**, 502-505.

GHT Gibbons, G. W., Hawking, S. W. and Townsend, P. K., editors (1986), *Supersymmetry and its Applications: Superstrings, Anomalies and Supergravity*, (Cambridge University Press).

GO Goddard, P. and Olive, D. (1984), Algebras, Lattices and Strings, in *Vertex Operators in Mathematics and Physics*, (Springer-Verlag) pp. 51-96.

GOR Gorenstein, Daniel (1982), *Finite Simple Groups: An Introduction to Their Classification*, (Plenum Press, New York).

GSW Green, M. B., Schwartz, J. H. and Witten, E. (1987), *Superstring Theory*, Vol. 1 and 2, (Cambridge University Press).

HUN Hungerford, Thomas W. (1974), *Algebra*, (Springer-Verlag, New York).

KLT Kawai, Hikaru, Lewellen, David C. and Tye, S.-H. Henry (1987), Construction of Fermionic String Models in Four Dimensions, *Nucl. Phys.* **288B**, 1-76.

KNE Kneser, M. (1967), Lineare Relationen zwischen Darstellungsanzahlen quadratisher Formen, *Math. Ann.* **168**, 31-39.

LAM1 Lam, C. S. (1988), GSO Projections of Modular-Invariant Aperiodic Strings, *Int. J. Mod. Phys.* **3A**, 913-942.

LAM2 Lam, C. S. and Li, Da-Xi (1985), Direct Compactification of Heterotic Strings From 26 to 4 Dimensions via T^{22}/G, *Commun. Theor. Phys.* **5**, 113-122.

LAM3 Lam, C. S., Strings Constructed From Free Bose and Fermi Fields, in *Proceedings of Beijing String Workshop*, (World Science).

LAM4 Lam, C. S. (1989), (in preparation).

NIE Niemeier, H.-V. (1973), Definite quadratische Formen der Dimension 24 und Diskriminante 1, *J. Number Theory* **5**, 142-178.

PIS Pissanetzky, Sergio (1984), *Sparse Matrix Technology*, (Academic Press).

SER Serre, J.-P. (1973), *A Course in Arithmetic*, (Springer-Verlag).

WAL Waldrop, M. Mitchell (Sept., 1985), Strings as a Theory of Everything, *Science* **229**, 1251-1253.

WIT Witt, E. (1941), Spiegelungsgruppen und Aufzählung halbeinfacher Liescher Ringe, *Abhandlungen Math. Sem. Univ. Hambury* **14**, 289-322.