Datagraphs in Algebraic Geometry and K3 Surfaces

Gavin Brown

Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK
gavinb@maths.warwick.ac.uk

Abstract. *Datagraphs* are combinatorial graphs having database items at their vertices and geometric relationships along their edges. I describe their applicability to lists of examples in algebraic geometry generated by computer algebra, and illustrate this with a list of K3 surfaces as the database items. The main point is that when analysing a single surface during construction of the database, the datagraph makes available its close relations, and that very often these provide extra information required to complete the analysis.

1 Introduction: lists in algebraic geometry

Beautiful lists of examples pervade algebraic geometry. Sometimes they comprise genuine classifications, such as the list of Du Val singularities in Table 1, see [D] for instance. At other times they are simply a list of known examples

(A_n)	$x^2+y^2+z^{n+1}=0$	for $n \geq 1$
(D_n)	$x^2+y^2z+z^{n-1}=0$	for $n \geq 4$
(E_6)	$x^2+y^3+z^4=0$	
(E_7)	$x^2+y^3+yz^3=0$	
(E_8)	$x^2+y^3+z^5=0$	

Table 1. Du Val surface singularities

making no pretence of being comprehensive. In either case one expects there to be relationships between items on the list. A datagraph is a device to record geometric objects together with these relationships as separate pieces of data supporting one another. There are a lot of mathematical issues, but here I focus on the symbolic calculations that make the general method work.

Perhaps the best list of all is that which constitutes the statement of the classification of surfaces: if S is a nonsingular projective algebraic surface, defined over $\mathbb{C}$ say, then S is birational to a nonsingular surface $\widetilde{S}$, a *minimal model* of S, in exactly one of the classes listed in Table 2. The Kodaira dimension κ is the order of growth of sections of nK_S with n, that is $h^0(nK_S) \sim n^\kappa$ with a conventional $\kappa = -\infty$ when no multiple has any sections. See Beauville [B] or Reid [R2] for an account of this classification.

κ	class of minimal surfaces
2	surfaces of general type
1	certain elliptic fibrations
0	Abelian, K3, Enriques or bielliptic surfaces
$-\infty$	ruled or rational surfaces

Table 2. The classification of surfaces

The key example studied here is a list of K3 surfaces, described in detail below. But the general methods used can be applied to many similar situations in algebraic geometry. Certainly many aspects of each class of surface in Table 2 have appeared in lists. For example, Reid [R1] worked out the case of a particular type of surface of general type. In fact, the approach described here has evolved from Reid's work over 20 years starting with this calculation. Lists of subcanonical curves and Enriques surfaces are in progress. Szendrői [Sz] has lists of Calabi–Yau 3-folds, three-dimensional analogues of K3 surfaces, and physicists have many more [KS]. And there are plenty of other lists of K3 surfaces from different points of view, the extremal elliptic fibrations of Shimada and Zhang [SZ] for instance, or [KS] again.

As well as being enjoyable in their own right, these lists are very useful. Belcastro [BC] computes the Picard lattices of a list of K3 hypersurfaces known as the famous 95, described below, in a study related to mirror symmetry. Corti, Pukhlikov and Reid [CPR] use the same famous 95 as a hunting ground for examples of *birational links*, the basic birational maps between 3-folds. Although their results are for 3-folds, much of the calculation occurs on a hyperplane section, which in this case is a K3 surface.

Let me clarify the main point. The aim is to construct databases, exploiting relationships between objects of the final database during the construction. First one accumulates a lot of plausible but rather raw data that is implicit in the problem: collections of singularities that are expected to appear on the objects is a typical example. Then refinement steps apply geometrical constructions to the raw data to construct the desired objects, or perhaps to prove that they don't exist. If X and Y appear in the database and we have a map $X \to Y$ between them, then the construction of X may be able to inherit calculations used in the construction of Y. In fact, this is exactly what happens in the K3 case discussed here. Section 4 below illustrates this point with a naive example, while Section 5 works out the K3 case. In the final section we glimpse the Magma [Ma] implementation, at least enough to see how it follows the model.

As an application, the main theorem, discussed in Section 2, uses a trivial computer analysis of the datagraphs to identify K3s that are 'primitive' from the point of view of projection. It is worth remarking that although many computer-generated lists exist in the geometry literature, I cannot think of a single one that is available in a format to be analysed seriously using a computer, rendering them pointless if they are at all large.

2 K3 surfaces and polarisation

A *K3 surface* is a simply-connected projective surface having only A_n singularities and trivial canonical class. (This is a slight modification of the usual definition, but it suits the discussion here and does no harm.) If S is a K3 surface in this sense, then the minimal resolution of singularities $\widetilde{S}$ of S appears in the list of Table 2. But the definition of K3 is not so important here: we don't use it directly at all. It appears only implicitly in the Riemann–Roch formula (1) of Section 5. A favourite example of K3 is the Kummer surface of a curve of genus 2, a surface of degree 4 in $\mathbb{P}^3$ with sixteen A_1 singularities coming from the fixed points of the involution -1 on the Jacobian.

In 1979, Reid wrote down the complete list of 95 weighted hypersurfaces that are K3 surfaces, the 'famous 95' (see them in [Fl] or [K3]). An example is

$$X_{30} \subset \mathbb{P}(5,6,8,11)$$

that is, the locus of zeros of a general homogeneous polynomial of degree 30 in variables x, y, z, t, which themselves have weights $5, 6, 8, 11$ respectively. After a change of coordinates, any such surface can be written as

$$f_{30} = 0 \quad \text{where} \quad f_{30} = x^6 + y^5 + yz^3 + zt^2 + \lambda xyzt$$

for some $\lambda \in \mathbb{C}$. Notice then that although one refers to *the* K3 surface X_{30}, it actually denotes a whole family of K3 surfaces parametrised by λ. Implicitly one usually has in mind only the general elements of this family, although special members of the family will be interesting too.

Notice that such K3 surfaces are automatically *polarised*. That is, they come as a pair X, A of a K3 surface X together with an ample Weil divisor A. The class A contains the divisors of rational functions of weight 1. Consider the example X_{30}. The divisor of y/x lies in the class of A but it is clearly not an effective divisor. In fact, the smallest monomial weight is 5 so $H^0(nA) = 0$ for $n = 1, 2, 3, 4$ while $H^0(5A)$ is 1-dimensional with basis x. Nonetheless, A is ample since multiples of it really do embed X and we still refer to A as the *hyperplane section*, even though there is no effective divisor in its class.

One calculates that the general X_{30} has singularities A_1, A_7 and A_{10}. These singularities are also polarised. For example, $\mathrm{A}_7 \in X_{30}$ is the singularity at the point $p_z = (0{:}0{:}1{:}0)$. Since $z = 1$ near p_z, f_{30} describes y as an implicit function and we can eliminate y there, realising the A_7 singularity as a $\mathbb{Z}/8$ quotient of $\mathbb{C}^2$. Such a quotient can happen in several ways. In this case, the global functions x, t restrict to the local axes u, v of the quotient and, modulo 8, the weights of x, t are $5, 3$. This polarised singularity is often denoted $\frac{1}{8}(3,5)$. As a reasonable shorthand for computer input we denote such singuarities by a pair of coprime integers $[r, a]$ with $1 \le a \le r/2$ where r is the order of the cyclic group, called the *index* of the polarised singularity, and a is the smaller of its weights. So the A_7 singularity on X_{30} is denoted $[8, 3]$. Fletcher's paper [Fl] contains a tutorial on these calculations.

One can project away from either of the singularities A_7 and A_{10}. After projection from A_7 the resulting surface is again a K3 surface, this time

$$X_{66} \subset \mathbb{P}(5, 6, 22, 33).$$

Such a projection is called a *Gorenstein projection*. The definition is rather involved so suffice it to say that a Gorenstein projection from a singularity $\frac{1}{r}(a, b)$ where $a + b = r$ results in a surface with two singularities $\frac{1}{a}(r, -r)$ and $\frac{1}{b}(r, -r)$. (If $a = 1$ then the $\frac{1}{a}$ point is nonsingular so we omit it; similarly for b.) The singularities of X_{66} are $\frac{1}{2}(1, 1)$, $\frac{1}{3}(1, 2)$, $\frac{1}{5}(2, 3)$ and $\frac{1}{11}(5, 6)$ so this projection calculus works, breaking up the $\frac{1}{8}(3, 5)$ into two singularities.

It is often the case that Gorenstein projection is simply elimination of a single variable, as is the case with usual projection. This is called Type I Gorenstein projection. It occurs when the local polarising weights a, b are also the weights of global variables, that is, when a, b appear among the weights of the ambient weighted projective space. The A_{10} singularity above is actually $\frac{1}{11}(5, 6) \in X_{30}$ — the monomial zt^2 eliminates z at this point — so projection from there is Type I, whereas projection from $\frac{1}{8}(3, 5) \in X_{30}$ is not. Projections of other types are also defined.

Since the famous 95 were computed, a number of Reid's students have extended the list to higher codimension: Fletcher [Fl] found 84 codimension 2 cases and Altınok [A] found more in codimension 3 and 4. In total these lists contain 391 K3 surfaces. (A tentative list of 162 codimension 5 K3s is now at [K3].)

There are interesting questions of verification for these lists. Reid's list of hypersurfaces is known to be complete — for instance, Johnson and Kollár [JK] have a proof (and a list of several thousand Fano 4-folds). This means that if $X \subset \mathbb{P}(a, b, c, d)$ is a K3 hypersurface then it is one of the famous 95 (although not necessarily a general element of the family). In higher codimension there is no such proof, although the lists have survived for several years and are probably correct if one rules out certain degenerations which were deliberately overlooked. A more immediate problem is to prove that every element on these lists actually exists. This is easy for the famous 95 since one can write down equations and compute. Similar explicit calculations work in codimensions 2 and 3. But in codimension 4 one has no structure theorem explaining how to write down the equations. So Altınok [A] resorts to a clever trick, called *unprojection*, to justify much of her list. The idea is that if a variety exists one can make a Gorenstein projection from one of its singularities, as in the example $X_{30} \dashrightarrow X_{66}$ above. With luck the result will be a K3 surface in lower codimension. If one can identify that K3 in advance and make certain polynomial calculations it is sometimes possible to reverse the projection process and view it as a construction. Thus while some items on Altınok's lists are not described as explicitly as X_{30}, they certainly all exist and carry implicit methods for their construction.

I will describe a model for constructing such lists as part of a computer algebra package. This model has been realised to build lists of K3 surfaces in the Magma system [Ma]. The current version 2.9 of Magma contains the database of 391 known K3 surfaces in codimension at most 4. I will discuss a substantial revision of this which already exists in prototype. (See [K3] for the most recent information.) Of course one benefits at every step from extra mathematical knowledge, but even if that is hard to come by one can make progress. This is typified by the amazing examples of [BR] which were found even before we knew about the unprojections involved. Or consider the K3 datagraphs themselves. Anyone who knows about K3 surfaces will have noticed that there is no mention of Picard lattices, one of the basic pieces in the theory of K3 surfaces. Indeed they play no role in the initial construction of the database, although it would be possible to use them. Subject specific details like these can be applied to basic datagraphs at a later stage.

As an application of the K3 datagraphs I prove the following theorem.

Theorem 1. *Let X, A be a polarised K3 surface having Du Val singularities and no Gorenstein projection to another polarised K3 surface. Let $g = \dim H^0(A) - 1$ and assume in addition that $g \le 2$. Then X, A has the same Hilbert series as one of the following:*

g	X
-1	$\operatorname{codim} X \le 4$ *and X is one of the 36 cases in Table 4*
0	$\operatorname{codim} X = 1$ *and X is one of the 6 cases in Table 3*
1	$X = X_{12} \subset \mathbb{P}(1,1,4,6)$
2	$X = X_6 \subset \mathbb{P}(1,1,1,3)$

In Tables 3 and 4, each K3 surface X is identified by its number in the K3 datagraph of the same genus: $\frac{1}{r}$ is shorthand for $\frac{1}{r}(1, r-1)$ and equation (2) in Section 5 determines A^2. More detailed Magma output is at [K3].

No.	K3	Basket $\mathcal{B}$	Degree A^2
1	$X_{42} \subset \mathbb{P}(1,6,14,21)$	$\frac{1}{2}, \frac{1}{3}, \frac{1}{7}$	1/42
12	$X_{36} \subset \mathbb{P}(1,5,12,18)$	$\frac{1}{5}(2,3), \frac{1}{6}$	1/30
23	$X_{30} \subset \mathbb{P}(1,4,10,15)$	$\frac{1}{2}, \frac{1}{4}, \frac{1}{5}$	1/20
77	$X_{24} \subset \mathbb{P}(1,3,8,12)$	$2 \times \frac{1}{3}, \frac{1}{4}$	1/12
249	$X_{18} \subset \mathbb{P}(1,2,6,9)$	$3 \times \frac{1}{2}, \frac{1}{3}$	1/6
794	$X_{10} \subset \mathbb{P}(1,2,2,5)$	$5 \times \frac{1}{2}$	1/2

Table 3. K3 surfaces with $g = 0$ and no Gorenstein projections

What is this theorem about? If a general K3 surface X, A has $g = 3$ then the sections of A embed X (modulo -2 curves) as a hypersurface $X_4 \subset \mathbb{P}^3$. The generality assumption in this case is exactly the same as asking that a genus 3 curve be nonhyperelliptic: indeed, X, A is general in this sense if and only if the general element of the linear system of A is a nonhyperelliptic curve of genus 3. If $g > 3$ then again with some generality hypothesis A already embeds X in an ordinary projective space. One could view these $g \ge 3$ cases as satisfactory. The theorem studies the other cases, those for which A does *not* have enough sections to embed X, assuming in addition that X cannot be constructed by

No.	K3	Basket $\mathcal{B}$	Degree A^2	cod
1	$X_{24,30} \subset \mathbb{P}(8,9,10,12,15)$	$\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}(2,3), \frac{1}{9}$	1/180	2
4	$X_{50} \subset \mathbb{P}(7,8,10,25)$	$\frac{1}{2}, \frac{1}{5}(2,3), \frac{1}{7}(3,4), \frac{1}{8}$	1/280	1
5	$X_{36} \subset \mathbb{P}(7,8,9,12)$	$\frac{1}{3}, \frac{1}{4}, \frac{1}{7}(2,5), \frac{1}{8}$	1/168	1
9	$X_{18,30} \subset \mathbb{P}(6,8,9,10,15)$	$2 \times \frac{1}{2}, 2 \times \frac{1}{3}, \frac{1}{5}, \frac{1}{8}$	1/120	2
12	$X \subset \mathbb{P}(6,6,7,8,9,10,11)$	$2 \times \frac{1}{2}, 2 \times \frac{1}{3}, \frac{1}{6}, \frac{1}{7}$	1/42	4
18	$X_{40} \subset \mathbb{P}(5,7,8,20)$	$\frac{1}{4}, 2 \times \frac{1}{5}(2,3), \frac{1}{7}$	1/140	1
21	$X_{66} \subset \mathbb{P}(5,6,22,33)$	$\frac{1}{2}, \frac{1}{3}, \frac{1}{5}(2,3), \frac{1}{11}(5,6)$	1/330	1
22	$X_{38} \subset \mathbb{P}(5,6,8,19)$	$\frac{1}{2}, \frac{1}{5}, \frac{1}{6}, \frac{1}{8}(3,5)$	1/120	1
24	$X_{27} \subset \mathbb{P}(5,6,7,9)$	$\frac{1}{3}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}(2,5)$	1/70	1
26	$X_{16,\ldots,20} \subset \mathbb{P}(5,6,7,8,9,10)$	$\frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \frac{1}{5}(2,3), \frac{1}{7}$	1/42	3
33	$X \subset \mathbb{P}(5,6,6,7,8,9,10)$	$\frac{1}{2}, \frac{1}{3}, \frac{1}{5}(2,3), 2 \times \frac{1}{6}$	1/30	4
42	$X_{14,\ldots,18} \subset \mathbb{P}(5,5,6,7,8,9)$	$\frac{1}{5}, 2 \times \frac{1}{5}(2,3), \frac{1}{6}$	1/30	3
48	$X_{34} \subset \mathbb{P}(4,6,7,17)$	$2 \times \frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{7}(2,5)$	1/84	1
50	$X_{16,18} \subset \mathbb{P}(4,6,7,8,9)$	$2 \times \frac{1}{2}, \frac{1}{3}, 2 \times \frac{1}{4}, \frac{1}{7}$	1/42	2
59	$X_{54} \subset \mathbb{P}(4,5,18,27)$	$\frac{1}{2}, \frac{1}{4}, \frac{1}{5}(2,3), \frac{1}{9}(4,5)$	1/180	1
61	$X_{32} \subset \mathbb{P}(4,5,7,16)$	$2 \times \frac{1}{4}, \frac{1}{5}, \frac{1}{7}(2,5)$	1/70	1
64	$X_{30} \subset \mathbb{P}(4,5,6,15)$	$2 \times \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, 2 \times \frac{1}{5}$	1/60	1
68	$X_{14,16} \subset \mathbb{P}(4,5,6,7,8)$	$\frac{1}{2}, 2 \times \frac{1}{4}, \frac{1}{5}(2,3), \frac{1}{6}$	1/30	2
130	$X_{12,14} \subset \mathbb{P}(4,4,5,6,7)$	$2 \times \frac{1}{2}, 3 \times \frac{1}{4}, \frac{1}{5}$	1/20	2
162	$X_{24} \subset \mathbb{P}(3,6,7,8)$	$\frac{1}{2}, 4 \times \frac{1}{3}, \frac{1}{7}$	1/42	1
167	$X_{48} \subset \mathbb{P}(3,5,16,24)$	$2 \times \frac{1}{3}, \frac{1}{5}, \frac{1}{8}(3,5)$	1/120	1
170	$X_{21} \subset \mathbb{P}(3,5,6,7)$	$3 \times \frac{1}{3}, \frac{1}{5}(2,3), \frac{1}{6}$	1/30	1
184	$X_{42} \subset \mathbb{P}(3,4,14,21)$	$\frac{1}{2}, 2 \times \frac{1}{3}, \frac{1}{4}, \frac{1}{7}(3,4)$	1/84	1
191	$X_{24} \subset \mathbb{P}(3,4,5,12)$	$2 \times \frac{1}{3}, 2 \times \frac{1}{4}, \frac{1}{5}(2,3)$	1/30	1
200	$X_{18} \subset \mathbb{P}(3,4,5,6)$	$\frac{1}{2}, 3 \times \frac{1}{3}, \frac{1}{4}, \frac{1}{5}$	1/20	1
269	$X_{10,12} \subset \mathbb{P}(3,4,4,5,6)$	$\frac{1}{2}, 2 \times \frac{1}{3}, 3 \times \frac{1}{4}$	1/12	2
420	$X_{15} \subset \mathbb{P}(3,3,4,5)$	$5 \times \frac{1}{3}, \frac{1}{4}$	1/12	1
544	$X_{30} \subset \mathbb{P}(2,6,7,15)$	$5 \times \frac{1}{2}, \frac{1}{3}, \frac{1}{7}$	1/42	1
551	$X_{42} \subset \mathbb{P}(2,5,14,21)$	$3 \times \frac{1}{2}, \frac{1}{5}, \frac{1}{7}(2,5)$	1/70	1
555	$X_{26} \subset \mathbb{P}(2,5,6,13)$	$4 \times \frac{1}{2}, \frac{1}{5}(2,3), \frac{1}{6}$	1/30	1
580	$X_{22} \subset \mathbb{P}(2,4,5,11)$	$5 \times \frac{1}{2}, \frac{1}{4}, \frac{1}{5}$	1/20	1
608	$X_{30} \subset \mathbb{P}(2,3,10,15)$	$3 \times \frac{1}{2}, 2 \times \frac{1}{3}, \frac{1}{5}(2,3)$	1/30	1
627	$X_{18} \subset \mathbb{P}(2,3,4,9)$	$4 \times \frac{1}{2}, 2 \times \frac{1}{3}, \frac{1}{4}$	1/12	1
820	$X_{12} \subset \mathbb{P}(2,3,3,4)$	$3 \times \frac{1}{2}, 4 \times \frac{1}{3}$	1/6	1
1611	$X_{14} \subset \mathbb{P}(2,2,3,7)$	$7 \times \frac{1}{2}, \frac{1}{3}$	1/6	1
2577	$X_{6,6} \subset \mathbb{P}(2,2,2,3,3)$	$9 \times \frac{1}{2}$	1/2	2

Table 4. K3 surfaces with $g = -1$ and no Gorenstein projections

unprojection from some other K3 surface. To avoid the question of generality, the theorem is phrased in terms of Hilbert series, as explained in Section 3. The (families of) K3 surfaces listed in the theorem each comprise an open subset of their Hilbert scheme, so X itself will lie in these families if it is sufficiently general.

Let me be clear about the proof of this theorem: one simply reads the result from the datagraphs of K3 surfaces for each $g \leq 2$ — I use a computer routine, but it could be done by hand. It could be viewed as fortunate that the result does not involve high codimension K3 surfaces that I would not be able to analyse, although the two codimension 4 surfaces in Table 4 — Altınok's two most difficult examples — are already mysterious. In Section 6, I indicate how the K3 datagraph is used in this proof: it could even be used to help to find a construction of these two surfaces by locating useful projections.

3 Varieties, graded rings and Hilbert series

I outline briefly the geometric background from [ABR]. Let R be a finitely generated graded ring. We usually assume that $R_0 = \mathbb{C}$ and that $R = \bigoplus_{n\geq 0} R_n$ with $R_n \cdot R_m \subset R_{n+m}$ and each R_n a finite dimensional vector space over $\mathbb{C}$. The *Hilbert series of* R is the formal power series

$$P_R(t) = \sum_{n=0}^{\infty} (\dim R_n)\, t^n.$$

Let X, A be a polarised variety, that is a projective variety X together with an ample Weil divisor A. (In practice both will be required to satisfy additional hypotheses.) It is common to take some very ample multiple nA of A and consider X in its embedding by the sections $H^0(nA)$ of nA. The classical example is when X is a nonhyperelliptic curve and A is its canonical class. Then A is already very ample and the embedding is the canonical embedding of X. Even if X is hyperelliptic, one can take $2A$ or $3A$ to embed it.

It is often better not to restrict to a single multiple of A, but to take them all at once. To do this, one embeds X by all generators of the finitely generated ring

$$R(X, A) = \bigoplus_{n\geq 0} H^0(nA)$$

in some *weighted* projective space. Then $X = \mathrm{Proj} R(X, A)$ and A is the natural $\mathcal{O}_X(1)$. Again there is a classical example, the embedding of a hyperelliptic curve $(y^2 = f(x)) \subset \mathbb{P}(1, g+1, 1)$ by all multiples of $A = 2p$ where $p \in X$ is a Weierstraß point. The weights of the projective space are the same as the degrees of the generators of $R(X, A)$ (to fix up the homogeneity of the embedding). In this case, the two 1s are the degrees of the functions defining the pencil $g^1_2 = |A|$ while the $g+1$ is the degree of the first section of a multiple of A which is not a power of the g^1_2.

So there is a general setup as follows:

$$\begin{array}{ccccc} \text{polarised variety} & & \text{graded ring} & & \text{Hilbert series} \\ X, A & \longleftrightarrow & R(X, A) & \longrightarrow & P_R(t) \end{array}.$$

where the arrows are natural translations. In an attempt to construct varieties we will work from right to left. The obvious problem is that there is no translation in that direction. Nonetheless, we will prescribe some invariants — ones that appear in the Riemann–Roch formula — of varieties X, A and use Riemann–Roch to compute $h^0(nA) = \dim H^0(nA)$ for each n. We package all these values in a

Hilbert series $P(t)$. (It is true that we have to concern ourselves with the higher cohomology, but often we work with varieties having good vanishing results for cohomology.) The challenge is then to construct X, A from the Hilbert series. This is not an exact science, but many tricks are known, most of which can be applied mechanically to a datagraph.

4 Computer generated lists and datagraphs

The point of datagraphs comes from the experience of building lists in algebraic geometry. Here I explain what datagraphs are and give a simplistic example. In Sections 5 and 6 I will work through the main example of this paper and show how they work in practice.

4.1 Datagraphs

To build a list of polarised geometric objects one first classifies the possible invariants that appear in a Riemann–Roch formula associated to the geometry. Each such piece of input generates a Hilbert series. The second stage, and the main problem, is to realise this as the Hilbert series of a suitable geometric object or even just show that such an object exists.

One thinks of the first part of the problem as being an easy combinatorial step (even when it is not) and the second part as being a difficult geometrical step. In the case of K3 surfaces, although we may have a plausible-looking Hilbert series, we have no general reason why there should be a surface X, A whose coordinate ring has this Hilbert series. And even if there is such an X, it may be difficult for us to prove that or to construct it.

Rather than regarding a database as being a list, we take it to be a directed combinatorial graph, the *datagraph*. It comprises a collection of vertices and a collection of directed edges, or arrows, between them. The basic skeleton of such a graph will be built at the combinatorial stage. Its vertices are in bijection with Hilbert series, while the directed edges correspond to Gorenstein projections between the corresponding K3 surfaces if they exist. The work required to prove existence, or deduce any other property of K3 surfaces does not then happen surface by surface, but for the most part inherits information from other surfaces by pulling it back along directed edges.

4.2 A naive example of the method

Suppose I present you with a set V of n numbered points in the plane, no three of which are collinear, and ask for a list of the combinatorial graphs having the points of V as vertices and straight lines as edges. This is an active problem, not one which any fixed database you might already own will answer directly.

It is certainly easy to write down a list of all possible graphs L_n on n numbered vertices (having at most one edge between any two vertices). This is a raw pre-database. Some items on it may be realised as planar graphs and some may not, and it may be hard to decide which.

The graph with no edges and also graphs having exactly one edge can all be realised on vertices V. So we attempt an induction on the number of edges. We check whether a graph in L_k with k edges can be realised by removing an edge, answering the question for the resulting graph on $k-1$ edges and then trying to redraw the edge. So there are two steps. The first is to identify in L_{k-1} those graphs which arise by removal of a single edge from a graph in L_k. This is part of the combinatorial step. The second is an algorithm which checks whether or not a new edge can be added to a given graph. This is the geometrical step.

The combinatorial step can be solved right at the beginning once and for all. Rather than taking the initial database of graphs to be a list of graphs $\cup L_n$, take it to be a *directed* graph G_c. The vertices of G_c are the graphs. There is a separate edge from any graph to each of the graphs which result by removing an edge. If you can implement an algorithm that checks whether a particular proposed edge can be added to a planar graph with result another planar graph, then you can set it working on G_c starting from the unique terminal vertex, the graph with no edges, creating the full subgraph G of G_c of all planar graphs, as required. The graph G is then a *datagraph.*

This datagraph certainly solves the initial problem — it lists all planar straight-line graphs on vertices V. And it already solves plenty of other problems. It is interesting to find maximal planar graphs, and in this context they are simply the vertices of G which have no edges directed into them. (For K3 surfaces, the interesting question is the complementary one of finding minimal vertices: this is Theorem 1.) Moreover, there are plenty of modifications to the question for which G_c is already useful. If I change the position of one of the n points, you only need to repeat the second geometrical step to G_c with the new points as basic data. Or maybe I relax the condition that edges must be straight lines, in which case one needs to re-implement the edge-adding algorithm before running it through G_c.

Presumably this is not a terribly effective approach to databases of graphs: such databases are usually large so the graph G_c would be unwieldy. But this approach is very effective for lists occurring in birational geometry: every feature of the process just described has a substantial counterpart there.

5 Building a datagraph of K3 surfaces

We will construct a datagraph of pairs X, A where X is a K3 surface and A is an ample Weil divisor (which we assume generates the analytic local class group at each singular point of X). The edges of the datagraph will be Gorenstein projections from the A_n singularities of X to other K3 surfaces.

5.1 The combinatorial step

Each K3 surface X, A is characterised by two pieces of data. The first is an integer $g \geq -1$, called the *genus*, equal to $h^0(A) - 1$ as in Theorem 1. The second is a *basket* $\mathcal{B}$ of singularities. A basket is a sequence of the polarised A_{r-1} singularities of X, $\frac{1}{r}(a, r-a)$ denoted by a pair of coprime integers $[r, a]$ with $1 \leq a \leq r/2$ as in Section 2.

The Riemann–Roch formula for K3 surfaces then computes $h^0(nA)$ for every $n \geq 1$. For any pair $(g, \mathcal{B})$, we collect these computations in a Hilbert series $P(t) = \sum_{n=0}^{\infty} h^0(nA)t^n$:

$$P(t) = \frac{1+t}{1-t} + \frac{t(1+t)}{2(1-t)^3}A^2 - \sum_{[r,a]\in\mathcal{B}} \frac{1}{2r(1-t^r)} \sum_{i=1}^{r-1} \overline{bi}(r - \overline{bi})t^i \tag{1}$$

where b is the reciprocal of a modulo r and $\overline{c}$ denotes the smallest residue of c modulo r. This is completed by the formula

$$A^2 = 2g - 2 + \sum_{[r,a]\in\mathcal{B}} \frac{b(r-b)}{r}. \tag{2}$$

Both formulas are worked out in Altınok's paper [A2].

There is one standard condition coming from the geometry of K3 surfaces which the basket of any K3 surface must satisfy:

$$\sum_{[r,a]\in\mathcal{B}} (r-1) \leq 19.$$

The point is that resolving the A_n singularities produces a nonsingular K3 surface. Such a surface has at most 20 linearly independent divisors lying on it. One is the hyperplane section A. Each A_n singularity contributes a further $n-1$ curves all independent of one another and of A. Imposing only this condition (and K3 experts will be aware of others), it is easy to check that there are exactly 6640 possible baskets that any K3 surface might have. So the collection of all possible combinatorial input is the infinite collection of pairs $(g, \mathcal{B})$ where $g \geq -1$ is an integer and $\mathcal{B}$ is one of the 6640 possible baskets. Since A is ample, it must be true that $A^2 > 0$ which is expressed by equation (2) as a condition on the pair $(g, \mathcal{B})$. One checks that this condition reduces the number of allowable baskets as follows:

genus	number of baskets
-1	4293
0	6491
1	6639
≥ 2	6640

.

The simple tricks one uses to try to construct a variety from a Hilbert series are far more effective when working with low genus so restricting our attention to pint-sized g, less that 2 say, is the better part of valour for now.

The combinatorial solution G_c is a directed graph as follows. First fix a small genus g. The graph G_c will have vertices in bijection with the collection of all baskets $\mathcal{B}$ for which $A^2 > 0$. For each pair $(g, \mathcal{B})$ we compute the Hilbert series $h_{\mathcal{B}}$. These power series have a natural order: $h_{\mathcal{B}_1} < h_{\mathcal{B}_2}$ if and only if the leading coefficient of $h_{\mathcal{B}_2} - h_{\mathcal{B}_1}$ is positive. We number the vertices $v_1, v_2 \dots$ of G_c in this order.

We now define a surgery operation on baskets. Let $[r, a]$ be any singularity in a basket $\mathcal{B}$ and let $r = a + b$. Make a new basket by replacing $[r, a]$ with the two singularities $[a, b']$ and $[b, a']$, where b' is $b \mod a$ and a' is $a \mod b$. The notation is a bit unhelpful here: we must interpret each of these new singularities correctly. If $r = 2$ we simply delete $[r, a]$. If $a = 1$ then we do not include $[a, b']$ at all. If $b' > a/2$ then write $[a, a - b']$ rather than $[a, b']$; similarly for $[b, a']$.

There is an edge from vertex v_i to v_j if and only if $\mathcal{B}_j$ can be constructed from $\mathcal{B}_i$ by a single such basket surgery operation. Gorenstein projection is decreasing with respect to the order on Hilbert series so one only needs to search for images among v_j with $j < i$.

5.2 Geometrising the vertices

At this stage we have a directed graph G_c which is the skeleton of our datagraph. Its vertices correspond to plausible Hilbert series for the graded homogeneous coordinate rings of polarised K3 surfaces, although we have only checked two simple conditions on the series so we must expect that sometimes no such surface actually exists. Its edges correspond to surgery operations on baskets that will be realised by Gorenstein projections if the corresponding surfaces exist. Now we begin to turn this into a datagraph of K3 surfaces.

Suppose given a Hilbert series $P(t) = \sum_{n \geq 0} h^0(A) t^n$. This power series can be written as a rational function

$$P(t) = \frac{p(t)}{(1 - t^{d_1}) \cdots (1 - t^{d_r})} \tag{3}$$

where $p(t)$ is a polynomial in t and $d_1, \dots, d_r$ are positive integers. As described in [R3] Section 3.6, if X is embedded in $\mathbb{P}(d_1, \dots, d_r)$ with hyperplane section A and Hilbert series $P(t)$ then a minimal free resolution of the graded ring $R(X, A)$ determines a rational expression for $P(t)$ in the form of equation (3) with Hilbert numerator

$$p(t) = 1 - \sum t^{e_j} + \cdots \pm t^k \tag{4}$$

where the integers e_j are the weights of the equations of X and k is the adjunction number of X. (Unfortunately the syzygies between the equations will contribute positive terms in t, some of which might well mask some of the t^{e_j}. But that is not the immediate problem.) How does one find this expression in advance?

In fact it is rather easy to generate expressions like (3). For example, if $g = -1$ and the basket $\mathcal{B}$ comprises $[2, 1]$, $[8, 3]$, $[11, 5]$ then one sees that

$$P(t) = \frac{t^{30} - 1}{(1 - t^5)(1 - t^6)(1 - t^8)(1 - t^{11})}. \tag{5}$$

One way is to write down the rational form of equation (1) immediately and manipulate that, giving a rational expression for $P(t)$: one constructs the unique minimal expression by cancelling common factors from numerator and demoninator. To understand $P(t)$ as a variety in weighted projective space we want to see the form of equation (3). There are many such expressions. To make one, we simply need to ensure that there are enough factors $1-t^d$ in the denominator to absorb all the cyclotomic factors in the minimal denominator. When the target variety X, A is a complete intersection, there is a unique minimal choice of such d_i; in codimension 3 or more complete intersections are rare so we must expect things to be more complicated.

The $1-t^d$ factors in the denominator are there to account for the periodic contributions arising from the singularities so the index of each singularity must divide the d occuring in some factor. For X_{30} this gives a first guess for a denominator of $(1-t^8)(1-t^{11})$. The ring $R(X,A)$ is 3-dimensional so we need to include another factor before the numerator becomes a polynomial. The factor $1-t$ is harmless since it is a factor any $1-t^d$. We compute then that

$$(1-t)(1-t^8)(1-t^{11})P(t) = t^{20} - t^{19} + t^{15} - t^{13} + t^{10} - t^7 + t^5 - t + 1.$$

Reading this Hilbert numerator, we see that there is an equation in degree 1. There is no point having a generator and an equation in degree 1 so the factor $1-t$ isn't quite right. We increase this power of t. Repeating the same reasoning reveals $1-t^5$ as the first sensible choice. Multiplying out again shows that a generator is needed in degree 6 and we are done.

We seek a projective variety whose homogeneous coordinate ring has the Hilbert series (5). Guided by the rule of thumb (4) we consider $X_{30} \subset \mathbb{P}(5,6,8,11)$. Any such variety has the right Hilbert series and as in the introduction one checks that if the equation defining X_{30} is general, then X_{30} is indeed a K3 surface with singularities as encoded by $\mathcal{B}$. Success!

5.3 Improving the datagraph: unprojection

By this stage we have a datagraph with candidates for K3 surfaces at its vertices and we would like to determine which really exist and which do not. They are not all as clear-cut as X_{30} above. In codimension higher than 3 we cannot simply write down equations and calculate. We will use the edges of the datagraph to guide an alternative construction process called *unprojection*.

Consider a special case of the main theorem of [PR]:

Theorem 2. *If $Y \subset \mathbb{P} = \mathbb{P}(d_1, \ldots, d_n)$ is a quasi-smooth K3 surface and one of the coordinate lines $D = \mathbb{P}(a,b) \subset \mathbb{P}$ lies on Y then there is a K3 surface $X \subset \mathbb{P}(d_1, \ldots, d_n, r)$ containing a singularity p of index $r = a+b$ and type $\frac{1}{r}(a,b)$. Moreover, Y is the Type I Gorenstein projection of X from p.*

One views this theorem as an implicit construction of X from the pair $Y \supset D$, an *unprojection*. Most K3 surfaces admit a Type I projection, especially when $g \neq -1$, so this approach promises some success. But Theorem 2 has a condition that needs to be checked. It is easy to find $D = \mathbb{P}(a,b)$ as a coordinate line, but we must also find an element of the family Y which contains this line. This is a monomial condition on the equations of Y so is easy to check if we know their degrees. But there is tension: too many conditions and the surface Y may acquire bad singularities and not be K3.

With this in mind, suppose we have a pair of candidate K3 surfaces with a Gorenstein projection between them $X \dashrightarrow Y$ and suppose inductively that we have a description of Y embedded in some weighted projective space. Now if X actually exists, and if the projection is Type I from a singularity $\frac{1}{r}(a,b) \in X$ then there is a curve $D \subset Y$ as in Theorem 2 and the weights of X are same as those of Y together with a new variable s of weight r. The projection is then simply the elimination of the variable s from the homogeneous coordinate ring of X and a, b are also weights of Y. (The surfaces have one trick up their sleeve. It can happen that some generators of Y are expressed in terms of the others once s is included so that not all the weights of Y are visible on X. In practice this only happens in codimension 1. The projection from $\frac{1}{11}(5,6) \in X_{30}$ is of this tricky Type I.)

So immediately we can tell whether there is any hope of realising X as a Type I unprojection of Y: the confirmed weights of Y should include a, b and the Hilbert series of X must be compatible with the weights of Y together with $a + b$, in the sense that an expression of the form (3) must be possible with these weights. When these conditions are satisfied we say that the projection $X \dashrightarrow Y$ is a *numerical Type I* projection. We also record the predicted weights of X. The task of proving that suitable Y exists remains — it is essentially a Gröbner basis problem, so something symbolic which could also be run through the datagraph — but already we expect to have a better collection of candidates. The experience of an analogous calculation in [BR] is that suitable Y often exist, but may do so in several distinct ways.

With datagraphs in hand, it is easy to write code to carry out this analysis. For example, one sees that

Proposition 1. *If $g = 1$ then there are unique weights for each of the 6639 possible K3 baskets so that each of the 20721 projections are either known Type I or numerical Type I projections with the single exception:*

$$\mathbb{P}(1,1,3,5) \supset X_{10} \dashrightarrow X_{12} \subset \mathbb{P}(1,1,4,6),$$

one of two projections to the $g = 1$ surface of Theorem 1.

In fact, all but six of the Type I projections are simply the elimination of a single variable. Apart from the exceptional case of the proposition, there are five other projections from K3s in codimension 1, which are necessarily the slightly tricky Type I. And let me emphasise again: there is no prima facie reason why any of these 6639 pieces of data should be realised by K3 surfaces, even though I expect most of them will be.

6 The Magma implementation

Datagraphs for K3 surfaces are implemented in prototype on Magma, see [K3] or [Ma] for information. A version that can be tailored to other geometric situations will be made in the future. Here we build the datagraph for K3 surfaces of genus -1 from scratch and then find the 36 surfaces of Table 4. The Magma commentary describes the steps discussed above. The > sign is the Magma prompt: the line following that is input, the rest is output. I leave you to work out the Magma syntax for yourself.

```
> G := K3BuildDatabaseGraph(-1);
  There are 4293 baskets with genus -1 ...       Time 108s
  Computing RR for baskets ...                   Time 173s
  Sorting the Hilbert series ...                 Time 229s
  Building a graph from the numerical data ...   Time 496s
  Analysing vertices of the graph ...            Time 830s
  Datagraph complete.
> K3PrintDatabase(G);
Database graph of 4293 K3 surfaces and 13078 projections
> #[ X : X in K3Surfaces(G) | #K3Projections(X) eq 0 ];
36
```

And that's it! The K3 surfaces that have no projection are exactly those on the datagraph with no arrows pointing away from them. And the number of these is exactly what the last line of input requested: remove the # to get the sequence of them all. Precisely, the result has identified Hilbert series of polarised K3 surfaces which cannot have a Gorenstein projection.

References

[K3] Lists of K3 surfaces are available from www.maths.warwick.ac.uk/~gavinb/k3graph.html and k3db.html

[A] Altınok, S. (1998): Graded rings corresponding to polarised K3 surfaces and $\mathbb{Q}$-Fano 3-folds. Thesis, Univ. of Warwick 93 + vii pp.

[A2] S. Altınok, Hilbert series and applications to graded rings, to appear in International journal of Mathematics and Mathematical science

[ABR] S. Altınok, G. Brown and M. Reid, Fano 3-folds, K3 surfaces and graded rings, to appear in Contemp Math (SISTAG 2001 Proceedings, ed Berrick et al.)

[B] A. Beauville, Complex algebraic surfaces. Second edition. LMS Student Texts, 34. CUP, 1996. x+132 pp

[BC] S-M. Belcastro, Picard lattices of families of K3 surfaces, Comm. Algebra 30 (2002), no. 1, 61–82

[BR] G. Brown and M. Reid, Extreme K3 surfaces with high unprojections, in preparation

[CPR] A. Corti, A. Pukhlikov and M. Reid, Birationally rigid Fano hypersurfaces, in Explicit birational geometry of 3-folds, A. Corti and M. Reid (eds.), CUP 2000, 175–258

[D] A. Durfee, Fifteen characterisations of Du Val singularities, Enseign. Math. (2) 25 (1979), no. 1-2, 131–163

[Fl] A.R. Iano-Fletcher, Working with weighted complete intersections, in Explicit birational geometry of 3-folds, CUP 2000, pp. 101–173

[JK] J. Johnson and J. Kollár, Fano hypersurfaces in weighted projective 4-spaces, Experiment. Math. 10 (2001), no. 1, 151–158

[KS] M. Kreuze and H. Skarke, Lists of K3 and Calabi-Yau on the web at hep.itp.tuwien.ac.at/∼kreuzer/CY/

[Ma] Magma (John Cannon's computer algebra system): W. Bosma, J. Cannon and C. Playoust, The Magma algebra system I: The user language, J. Symb. Comp. **24** (1997) 235–265.
See also www.maths.usyd.edu.au:8000/u/magma

[PR] S. Papadakis and M. Reid, Kustin–Miller unprojection without complexes, J. algebraic geometry (to appear), preprint math.AG/0011094, 15 pp

[R1] M. Reid, Surfaces with $p_g = 0$, $K^2 = 1$. J. Fac. Sci. Univ. Tokyo Sect. IA Math. 25 (1978), no. 1, 75–92

[R2] M. Reid, Chapters on algebraic surfaces. Complex algebraic geometry (Park City, UT, 1993), 3–159, IAS/Park City Math. Ser., 3, Amer. Math. Soc., Providence, RI, 1997

[R3] M. Reid, Graded rings and birational geometry, in Proc. of algebraic geometry symposium (Kinosaki, Oct 2000), K. Ohno (Ed.), 1–72, available from www.maths.warwick.ac.uk/∼miles/3folds

[SZ] I. Shimada and D-Q. Zhang, Classification of extremal elliptic $K3$ surfaces and fundamental groups of open $K3$ surfaces, Nagoya Math. J. 161 (2001), 23–54

[Sz] B. Szendrői, Calabi-Yau threefolds with a curve of singularities and counterexamples to the Torelli problem. Internat. J. Math. 11 (2000), no. 3, 449–459.