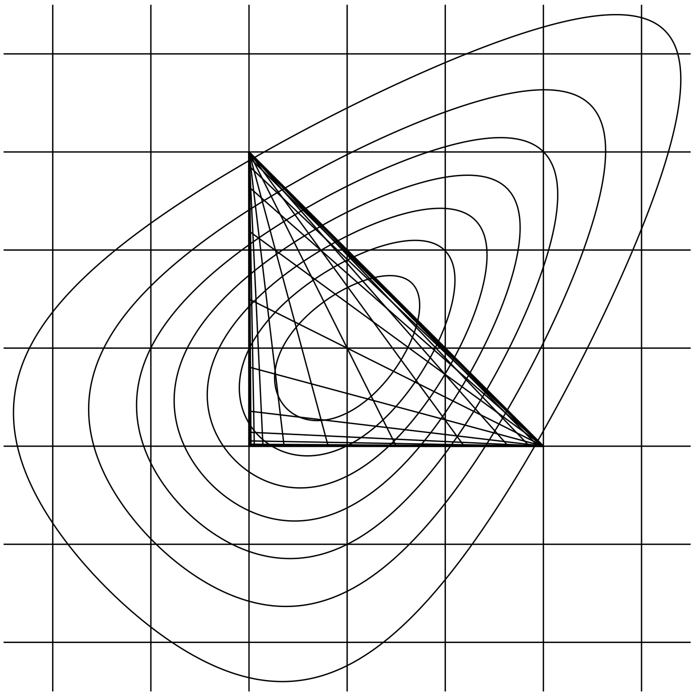
CHALMERS UNIVERSITY OF TECHNOLOGY | UNIVERSITY OF GOTHENBURG



Kähler-Einstein metrics on toric Fano manifolds and connections to Optimal Transport

Master's thesis in engineering mathematics and computational science

Rolf Andreasson

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2021

Master's thesis 2021

Kähler-Einstein metrics on toric Fano manifolds and connections to Optimal Transport

Rolf Andreasson

UNIVERSITY OF GOTHENBURG

CHALMERS UNIVERSITY OF TECHNOLOGY

Department of Mathematical Sciences
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2021

Kähler-Einstein metrics on toric Fano manifolds and connections to Optimal Transport

Rolf Andreasson

Examiner and Supervisor: Robert Berman, Department of Mathematical Sciences

Master's Thesis 2021
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: A visualization of the Kähler-Einstein measure of $\mathbb{C}\mathrm{P}^2$ with the associated fundamental polytope. The two grids shown show how the coordinate grid on $\mathbb{R}^2$ is transformed under the transport map.

Typeset in LaTeX
Gothenburg, Sweden 2021

Kähler-Einstein metrics on toric Fano manifolds and connections to Optimal Transport

Rolf Andreasson
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg

Abstract

We introduce the problem of finding a Kähler-Einstein metric on a Kähler manifold and specifically on a Fano manifold. We restrict to the class of toric complex manifolds where the symmetry can be used to reduce the resulting partial differential equation to a real equation in $\mathbb{R}^n$. We then introduce the theory of optimal transport, specially adapted to the application. We present a special transportation problem which is in fact equivalent to a weak formulation of the Kähler-Einstein equation on toric Fano manifolds. The presentation is a literature study aimed at presenting the material in a self-contained and elementary fashion.

We also present a novel variational approach to the existence problem in the language of optimal transport and equilibrium physics. We show some results towards an existence result based on this approach.

Finally we exemplify large parts of the theory on complex projective space, an explicit example of a toric Fano manifold. We also compute the free energy, an invariant we will introduce, on complex projective space.

Acknowledgements

I would like to thank my supervisor Robert Berman for continuous support and patience with my questions and for introducing me to this fascinating area of mathematics. Additionally, I would like to thank my friend and colleague Ludvig Svensson for his support and insightful comments.

Rolf Andreasson, Gothenburg, June 2021

Contents

1

Introduction

The motivation behind the study of Kähler–Einstein metrics is different if you ask different parts of the mathematical community. Firstly, there is the viewpoint of algebraic geometry. Algebraic geometry is the study of the common zero-set of one or more polynomials in several variables. If the underlying field is $\mathbb{C}$ then the resulting set admits the structure of a (possible singular) complex manifold. If the polynomials is homogeneous, i.e, $p(\lambda z) = \lambda^k p(z)$ where $\lambda \in \mathbb{C}$ and k is the degree, then we can consider the zero-set in projective space where we mod out the multiplication by a complex number. These spaces, which are now compact complex manifolds, are called projective complex algebraic varieties and have been heavily studied. On these spaces one can consider Riemannian metrics. The class of Kähler–metrics are these that are adapted to the complex and algebraic structure. Among these metrics one can look for canonical metrics. This leads naturally to Kähler–Einstein metrics.

From the viewpoint of differential geometry in large, a natural question since Gauss has been to study existence of canonical metrics. In fact, the question of whether a compact orientable smooth surface admits a constant curvature metric is a special case of the existence of Kähler–Einstein metrics on Riemann surfaces. In higher dimensions, one can consider constant-curvature metrics whose scalar curvature is constant, or the smaller class of Einstein metrics whose Ricci curvature is proportional to the metric itself. The study of these metrics is in general difficult but progress in the case when the metric is in addition Kähler has been considerable, especially in the case of Einstein metrics. Thus Kähler–Einstein metrics can be seen as a special, simpler case in this vast research area.

Thirdly there are deep connections to theoretical physics, in particular to supersymmetry. Supersymmetry arose as a concept in String theory, a popular quantum theory of gravity, but exists as a theoretical concept in general quantum field theory. It is a peculiar symmetry of certain field theories, in particular of ten-dimensional string theory. Since we observe a four-dimensional space time, six of those dimensions are conjectured to be compact and small. In vacuum solutions to the theory, demanding that at least one independent degree of freedom of the supersymmetry remains unbroken, the metric on the compactified space is forced to be Kähler–Einstein.

Within the area of Kähler–Einstein metrics itself there are important subareas. Recall the vacuum Einstein equation with cosmological constant,

$$\mathrm{Ric}(g) = \lambda g \tag{1.1}$$

where g is the metric and $\mathrm{Ric}(g)$ the Ricci curvature of g. After rescaling the metric there are essentially just three choices of λ: -1, 0 or 1. These correspond roughly to the three types of classical geometries, hyperbolic, flat and the geometry of the sphere, respectively. In fact the underlying space fixes the value of λ. These three different classes of spaces has names in the algebraic setting and they are varieties that are canonically polarized, Calabi-Yau and Fano varieties, respectively.

For $\lambda = -1$, existence of a unique Kähler–Einstein metrics was solved by Aubin [1] and Yau [2], and for the case $\lambda = 0$ by Yau [3]. The case $\lambda = 1$ is more subtle and there are obstructions to existence. Existence turns out to be equivalent to a certain stability condition on the underlying complex manifold called *polystability*. For spaces that are polystable there is a Kähler–Einstein metric unique up to holomorphic automorphisms, proven by Cheng–Donaldsson–Sun in [4]. The other direction was first proven by Berman in [5].

1.1 Outline

This thesis is concerned solely with the case $\lambda = 1$. It is also only concerned with the case when the underlying complex manifold is toric, a class of Kähler manifolds which significantly simplifies the analysis while still leaving it very rich. In this case, there is a surprising connection between Kähler–Einstein metrics and the theory of optimal transportation. The goal of the thesis is to introduce the subject of Kähler–Einstein metrics on toric Fano varieties in sections 2 and 3. Then introduce the theory of optimal transport in section 4. In the course of these sections the connection between the areas will be explained. Then in section 5, a variational formulation of Kähler–Einstein metrics on toric Fano manifolds is introduced.

While sections 2–4 can be considered a literature study section 5 is more exploratory as this particular approach is not present in the literature. Existence of Kähler–Einstein metrics on toric Fano manifolds has been successfully investigated via a variational problem before in [6] and also on general Fanos in [7]. Our approach here is akin to the variational approach in [7], but potentially avoiding the use of the energy-entropy compactness theorem and trying to simplify the analysis using the toric symmetry. The section contains some result towards an existence theory developed using this particular variational problem and a partial outline of what such a proof could look like.

Section 6 is devoted to the example of complex projective space, $\mathbb{C}\mathrm{P}^n$, the typical example of a toric Fano manifold where the theory presented in the thesis is exemplified and an explicit Kähler–Einstein metric can be found. Additionally, the free energy, an invariant we will define in section 5, is computed for complex projective space and is shown to, asymptotically in the dimension n, be smaller for $\mathbb{C}\mathrm{P}^n$ than for $(\mathbb{C}\mathrm{P}^1)^n$. This verifies in a particular case a conjecture of Berman that the free energy invariant is minimal for $\mathbb{C}\mathrm{P}^n$ among all Fanos. That is, it is verified asymptotically in the dimension n, in the specific case comparing $\mathbb{C}\mathrm{P}^n$ to $(\mathbb{C}\mathrm{P}^1)^n$.

Throughout the text we will assume knowledge of general differential geometry and basic measure theory.

2

Introductory Kähler geometry

Kähler geometry lies at the intersection of Complex, Symplectic and Riemannian Geometry. Our point of view will be mostly the complex one. In the introduction of complex geometry we will be rather brief and present the minimal amount of theory and leave any small proofs to the reader.

2.1 Complex geometry

Recall how in ordinary differential geometry, a smooth manifold is a space which locally look like $\mathbb{R}^n$, glued together in a smooth fashion so that smooth objects can be studied on them in a coordinate independent fashion. Complex geometry and complex manifolds is the answer to what we would need to study spaces that locally looks like $\mathbb{C}^n$ and on which we can study holomorphic objects.

Definition 1. *A* complex manifold *is a smooth manifold for which the local model is* $\mathbb{C}^n$ *and the transition maps are biholomorphic.*

Following the procedure from ordinary differential geometry one can define in a local fashion holomorphic functions and maps and check that these notions are coordinate independent and become global.

If we work locally in $\mathbb{C}^n$ with coordinates $z_k = x_k + iy_k$ then we get a split of the tangent space at the origin into

$$T_0(\mathbb{C}^n) = \operatorname{span}\left(\frac{\partial}{\partial x_k}\right) \oplus \operatorname{span}\left(\frac{\partial}{\partial y_k}\right). \tag{2.1}$$

We can define a map J which rotates a complex vector by i in each copy of the complex plane by

$$J\left(\frac{\partial}{\partial x_k}\right)=\frac{\partial}{\partial y_k} \tag{2.2}$$
$$J\left(\frac{\partial}{\partial y_k}\right)=-\frac{\partial}{\partial x_k} \tag{2.3}$$

We will be especially interested in the eigenvectors of this map and we define

$$\frac{\partial}{\partial z_k}:=\frac{1}{2}\left(\frac{\partial}{\partial x_k}-i\frac{\partial}{\partial y_k}\right) \tag{2.4}$$
$$\frac{\partial}{\partial \bar{z}_k}:=\frac{1}{2}\left(\frac{\partial}{\partial x_k}+i\frac{\partial}{\partial y_k}\right) \tag{2.5}$$

where these are now vectors in the complexified tangent space and it is easily checked that $J\left(\frac{\partial}{\partial z_k}\right)=i\frac{\partial}{\partial z_k}$ and $J\left(\frac{\partial}{\partial \bar{z}_k}\right)=-i\frac{\partial}{\partial \bar{z}_k}$. One can also check that due to the transitions maps being holomorphic, the split of the complexified tangent space into eigenspaces of J is invariant under change of coordinates and thus J is extended to a global map on the real or complexified tangent bundle of a complex manifold. J additionally satisfies $J^2=-\mathrm{Id}$ and in fact the concept of a complex manifold can be rephrased entirely in terms of J, called the complex structure. We can also introduce locally the one-forms

$$\mathrm{d}z_k=\mathrm{d}x_k+i\mathrm{d}y_k \tag{2.6}$$
$$\mathrm{d}\bar{z}_k=\mathrm{d}x_k-i\mathrm{d}y_k \tag{2.7}$$

and by linear algebra the the split of the tangent space yields a split of the space of degree N complex differential forms Ω^N splits into

$$\Omega^N=\bigoplus_{k+l=N}\Omega^{(k,l)}. \tag{2.8}$$

where $\Omega^{(k,l)}$ is the space of differential forms which locally only contains a linear combination of terms of the form $\mathrm{d}z_{i_1}\wedge\ldots\wedge\mathrm{d}z_{i_k}\wedge\mathrm{d}z_{j_1}\wedge\ldots\wedge\mathrm{d}z_{j_l}$. Denoting by $\pi^{(k,l)}$ the projection of a form onto $\Omega^{(k,l)}$ we have

$$d=\pi^{(k+1,l)}d+\pi^{(k,l+1)}d=:\partial+\bar{\partial}. \tag{2.9}$$

Locally for a function f,

$$\partial f = \sum_i \frac{\partial}{\partial z_i} f \mathrm{d}z_i \tag{2.10}$$

and this is the reason for the division by 2 in (2.5).

2.2 Kähler geometry

As was noted earlier Kähler geometry can be viewed from several perspectives, from the complex perspective, a Kähler manifold is a complex manifold equipped with a Kähler form ω.

Definition 2. *A Kähler form is a real, closed 2-form satisfying that $g = \omega(\cdot, J\cdot)$ is positive definite. Additionally ω is preserved by J i.e, $J_*\omega = \omega$.*

From this we immediately get that g is symmetric, indeed $\forall u, v$

$$g(u,v) = \omega(u, Jv) = -\omega(Jv, u) = -\omega(J^2 v, Ju) = g(v,u) \tag{2.11}$$

thus a Kähler manifold has a canonical metric. Since J is non-degenerate, ω will also be non-degenerate. A closed, non-degenerate 2-form is a precisely the definition of a symplectic form, making the underlying manifold into a symplectic manifold. Thus a Kähler manifold is simultaneously a Riemannian, complex and symplectic manifold and these three structures are connected via the Kähler condition above. We can split ω such that in in coordinates

$$\omega = \sum_{i,j} a_{ij} dz_i \wedge dz_j + b_{ij} dz_i \wedge d\bar{z}_j + c_{ij} d\bar{z}_i \wedge d\bar{z}_j. \tag{2.12}$$

Since ω is preserved under the action of J and $J^* dz = -idz$ we need $a_{ij} = c_{ij} = 0$. The positive definite condition now becomes that b_{ij}/i is positive definite. We will denote this type of positivity notion with $\omega > 0$.

2.3 Holomorphic line bundles

Our approach to Kähler metrics will be via hermitian metrics on holomorphic line bundles.

Definition 3. *Let X be a complex manifold. A holomorphic line bundle over X is a complex manifold L and a holomorphic mapping $\pi : L \to X$ such that for every point $p \in X$ there is a neighborhood U of p and a biholomorphic map $\psi : \pi^{-1}(U) \to U \times \mathbb{C}$ such that the following diagram commute*

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\psi} & U \times \mathbb{C} \\ \downarrow{\scriptstyle \pi} & \swarrow{\scriptstyle pr_1} & \\ U & & \end{array}$$

where pr_1 is the projection onto the first component. It is clear that pre-images of points (fibers) under π are biholomorphic to $\mathbb{C}$ and thus carry a vector space structure. ψ is additionally required to be linear on fibers with respect to this vector space structure. In that way the fibers can be identified with complex vector spaces in a trivialization invariant way.

The definition reflects the fact that while $X \times \mathbb{C}$ is always a holomorphic line bundle, called the trivial line bundle, a general holomorphic line bundle is only of such a form locally. A *holomorphic section* (which will mostly be referred to as just a section) of a line bundle is a holomorphic map $s : X \to L$ such that $\pi \circ s = \mathrm{Id}$, generalizing the concept of a holomorphic function which is a holomorphic section of the trivial line bundle. The holomorphic tangent and cotangent bundles are examples of higher dimensional versions of holomorphic line bundles although we will only encounter these concrete examples. Some terminology that will be used is that L in the above definition will be referred to as the total space and X as the base space.

There is also a one-to-one correspondence between trivializations and non-zero sections. Given a trivialization ψ as in the definition above, the map $p \mapsto \psi^{-1}(p, 1)$ gives a non-zero section of the line bundle $\pi^{-1}(U) \to U$. Conversely given a non-zero section e on U, we can write any $l \in L$ as $l = ze(\pi(l))$, for some $z \in \mathbb{C}$. Now, ψ defined by $l \mapsto (\pi(l), z)$ will be a trivialization of L over U. Additionally we note that, starting with a trivialization and proceeding as above, first constructing a non-zero section, and from it again a trivialization, we end up with the starting trivialization again. Thus the existence of a global non-zero holomorphic section is equivalent to the line bundle being trivial. When we refer to a trivialization we will refer sometimes to the biholomorphism and sometimes to the non-zero section depending on the context. The section will also be referred to as trivializing section.

Given two trivializing sections s_1 and s_2 of a vector bundle, with non-empty overlap, we can define the corresponding transition function. Indeed if $U, V \subset X$ such that $U \cap V \neq \emptyset$

then formally

$$\Psi_{12} = e_2/e_1 \tag{2.13}$$

is a holomorphic function locally defined that one can use to change trivialization between e_1 and e_2. In fact one can define the whole concept of line bundles only using transition functions instead of trivializations. The transition function essentially tells us how the local product spaces should be glued together, similar to the transition functions defining manifolds.

We will need a notion of when two line bundles on the same base space are considered isomorphic. Note that we can change trivialization on the same open cover $\{U_k\}_k$ by multiplying all trivializing sections $\{s_k\}$ by some non-zero holomorphic function $\{f_k\}$. Thus two line bundles such defined should clearly be considered isomorphic. For the transition maps this means that two sets of transition maps $\{\Psi_{kl}\}_{kl}$ and $\{\tilde{\Psi}_{kl}\}_{kl}$ are the same if we can find holomorphic functions $\{f_k\}$ on each open set in the cover such that

$$\tilde{\Psi}_{kl} = \frac{f_k}{f_l}\Psi_{kl}. \tag{2.14}$$

If two line bundles on the same base space X can be given transition maps on the same cover of X that satisfy relations of this kind then we will regard the lines bundles isomorphic.

We will consider metrics $||\cdot||$ on $L \to X$, varying smoothly over X. Given a trivialization e over $U \in X$, we can define what we will call the *weight function* $\phi = -\log(||e||^2)$ of $||\cdot||$. This function determines the metric over U since any other section s can be written $s = fe$ with f holomorphic and we get necessarily $||s|| = |f|e^{-\phi/2}$. However, ϕ will not be global and will generally depend on the choice of trivialization. Despite this, the following 2-form is globally well-defined.

Definition 4. *The curvature form ω of a metric $||\cdot||$ on a line bundle $L \to X$ is a global 2-form given locally by*

$$\omega = \frac{i}{2\pi}\partial\bar{\partial}\phi \tag{2.15}$$

where ϕ is the weight function of the metric $||\cdot||$ in any trivialization. [1]

[1] As the name suggests, this is the curvature form of a connection on the line bundle. The connection that has to be used is the unique so called *hermitian connection* which has a compatibility condition with the metric and another compatibility condition with the complex structure.

To see that this is well-defined let ψ_1, ψ_2 be two local trivializations and e_1 and e_2 be the corresponding non-zero sections over $U \subset X$. We have $\psi_1^{-1}\psi_2(e_2) = e_1$ and since this transition map is holomorphic and linear over the fibers we must have $fe_2 = e_1$ for f a non-zero holomorphic function on U. Evaluating the metric on e_1 with respect to the two trivialization we get

$$||e_1||^2 = e^{-\phi_1} = |f|^2 e^{-\phi_2}. \tag{2.16}$$

Thus

$$\partial\bar{\partial}\phi_1 = -\partial\bar{\partial}\log e^{-\phi_1} = -\partial\bar{\partial}\log|f|^2 e^{-\phi_2} = \partial\bar{\partial}\phi_2 - \partial\bar{\partial}\log|f|^2. \tag{2.17}$$

For the last term we get

$$\partial\bar{\partial}\log|f|^2 = \partial\bar{\partial}\log f + \partial\bar{\partial}\log\bar{f} = 0 \tag{2.18}$$

where $\bar{\partial}\log f = 0$ since f is locally holomorphic and $\partial\bar{\partial}\bar{f} = 0$ since $\bar{\partial}\log\bar{f}$ is locally antiholomorphic. Thus ω is well-defined as a global 2-form.

At this point it will be useful to note that if we introduce $d^c = \frac{-1}{2\pi}J^*d$ we have

$$dd^c = -\frac{1}{4\pi}(\partial+\bar{\partial})J^*(\partial+\bar{\partial}) = \frac{i}{4\pi}(\partial+\bar{\partial})(\partial-\bar{\partial}) = \frac{i}{4\pi}(\partial\bar{\partial}-\bar{\partial}\partial) = \frac{i}{2\pi}\partial\bar{\partial} \tag{2.19}$$

where we have again used $J^*(dz) = -idz$. Thus the curvature form ω is closed as locally we have $\omega = dd^c\phi$. Therefore we can talk about $[\omega]$, the cohomology class of ω. It is clear that the curvature form has some dependance on the metric but in fact the cohomology class does not.

Proposition 1. *The cohomology class $[\omega]$ is independent of the metric.*

Proof. Let $||\cdot||^{(1)}$ and $||\cdot||^{(2)}$ be two metrics on $L \to X$. In a local trivialization let $\phi^{(1)}$ and $\phi^{(2)}$ be the corresponding weight functions. We saw above that changing trivialization amounted to adding a strictly positive function to the weight function. Thus $g := \phi^{(2)} - \phi^{(1)}$ is a globally well-defined function. Now simply observe that

$$\omega^{(1)} = \frac{i}{2\pi}\partial\bar{\partial}\phi^{(1)} = \frac{i}{2\pi}\partial\bar{\partial}\phi^{(2)} + \partial\bar{\partial}\frac{i}{2\pi}g = \omega^{(2)} + dd^c g. \tag{2.20}$$

The difference $\omega^{(1)} - \omega^{(2)}$ is thus exact and we get $[\omega^{(1)}] = [\omega^{(2)}]$. □

This amusing fact fits in to the large theory of characteristic classes. For general vector bundles one can easily introduce natural morphisms between them that allows one to talk about a vector bundle modulo isomorphisms. Much as smooth manifolds and diffeomorphisms. One can then talk about invariants and try to classify vector bundles using these. Assigning cohomology classes to vector bundles is precisely this and in the case of holomorphic vector bundles, the class of the curvature form is the first Chern class of L, denoted $c_1(L)$.

Another thing we can wonder now that we know that the curvature form is closed, is whether it is in fact really a Kähler form. The remaining condition is the positivity condition. This is not always true, but we call those metrics whose curvature form is Kähler *positively curved.* One can wonder if all Kähler metrics are of this form. It is the case after restricting to integral cohomology classes but will not be necessary theory for this project.

The local weight ϕ of a positively curved metric satisfies then

$$i\partial\bar{\partial}\phi > 0 \tag{2.21}$$

in the sense that the local matrix of this 2-form is positive definite. A real smooth function satisfying this is called *strictly plurisubharmonic.* We already know that these functions might not generally be global, but if we start with a reference metric $||\cdot||_0$ with curvature form ω_0, we know that any other metric is given by $||\cdot|| = e^{-\phi/2}||\cdot||_0$ where ϕ is global. This metric is positively curved if and only if

$$i\partial\bar{\partial}\phi + \omega_0 > 0 \tag{2.22}$$

and we say that ϕ is ω_0- *strictly plurisubharmonic.* These classes of functions have very nice analytical properties that will help us understand the geometry so a bit of this theory is worthwhile developing.

Proposition 2. *Let ϕ be a strictly plurisubharmonic function on a complex manifold X. Then $\iota^*\phi$ is strictly plurisubharmonic on every complex curve on X (complex submanifold of dimension 1) where ι is the inclusion.*

Proof. Assume ϕ is strictly plurisubharmonic on X and let C be a complex submanifold of X. The statement is local and we introduce holomorphic coordinates z^j on X and w on C. Let F_i be the matrix representing ι^* locally with respect to the chosen holomorphic coordinates. Compute

$$i\partial\bar{\partial}\iota^*\phi = i\iota^*\partial\bar{\partial}\phi = i\frac{\partial^2\phi}{\partial z^i\partial\bar{z}^j}\iota^*(dz^i \wedge d\bar{z}^j) = i\frac{\partial^2\phi}{\partial z^i\partial\bar{z}^j}F^i\bar{F}^j dw \wedge d\bar{w} \tag{2.23}$$

where we have used in the first step that i^* commutes with $\partial\bar{\partial}$ which follows since $\partial\bar{\partial} = dd^c$ and d always commutes with pullbacks and since i is a holomorphic map by the definition of complex submanifold its pullback commutes with J^*. The scalar function in front of the differential forms in the last expression is positive by the definition of the positive definiteness of $i\frac{\partial^2\phi}{\partial z^i\partial\bar{z}^j}$. □

This allows us to prove this very nice property of strictly plurisubharmonic function:

Proposition 3. *There are no global strictly plurisubharmonic functions on a compact complex manifold.*

Proof. Let, for the case of contradiction, X be a compact complex manifold and ϕ a strictly plurisubharmonic function on X. Since X is compact and ϕ continuous ϕ attains its infimum over X at a point $x^* \in X$. Take now any complex submanifold C, not necessarily closed, of complex dimension 1 containing x^* and denote by I the inclusion of C into X. By Proposition 2 the function $I^*\phi$ is strictly plurisubharmonic on C. Concretely in holomorphic coordinates $w = x + iy$ on C

$$i\partial\bar{\partial}\phi = i\frac{\partial^2\phi}{\partial w\partial\bar{w}}dw \wedge d\bar{w} = \left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2}\right) idw \wedge d\bar{w}. \tag{2.24}$$

Consequently $\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2}\right) > 0$. Thus the trace of the hessian of ϕ at x^* is positive and x_* is not a local minimum of $I^*\phi$, thus not a global minimum of ϕ on X. □

We will use this to deduce the following classification of holomorphic line bundles with respect to their positivity properties.

Proposition 4. *Let X be a compact complex manifold and $L \to X$ a holomorphic line bundle on X. Then precisely one of the following holds*

- *L is positive in the sense that there is a positively curved metric on L.*

- *L is negative in the sense that there is a positively curved metric on L^*.*
- *L is the trivial line bundle.*
- *L is of indefinite sign and non of the above holds.*

Proof. We begin by showing that if L is positive then it is not negative. Assume $||\cdot||$ is a metric on L with positive curvature form ω_0. The point now is to consider the dual metric $||\cdot||^*$. The dual metric could just be defined to be the metric with weight function the negative of the weight of $||\cdot||$. One shows, by writing out the definitions, that this coincides with the standard definition of dual metric. The dual metric thus has has curvature form $-\omega_0$ and any other metric on L is determined by a function ϕ and has curvature form

$$\omega = \frac{i}{2\pi}\partial\bar{\partial}\phi - \omega_0. \tag{2.25}$$

Demanding the existence of a metric of positive curvature on L^* demands that we can find $\phi \in C^\infty(X)$ such that

$$\frac{i}{2\pi}\partial\bar{\partial}\phi > 0 \tag{2.26}$$

but by Proposition 2.3 there are no global strictly plurisubharmonic functions on a compact complex manifold. This ensures that L cannot be both positive and negative.

If L is the trivial line bundle we can define a metric by giving a global non-zero section constant length. The weight function for that metric is global and constant and its curvature is zero. Again any other metric can be represented by a global weight function ϕ and the curvature form will be

$$\omega = \frac{i}{2\pi}\partial\bar{\partial}\phi. \tag{2.27}$$

If the metric or its dual, represented by ϕ and $-\phi$, respectively, is to be positive, then again we get a global strictly plurisubharmonic function contradicting Proposition 2.3. Thus only one of the cases can be true simultaneously. The last statement takes care of the remaining cases.

□

The above proposition actually provides four categories not only of line bundles on compact complex manifolds, but gives four categories of complex manifolds themselves. This since there is a canonical line bundle on every complex manifold.

Definition 5. *Let X be a complex manifold of dimension n. Then we define the canonical bundle K_X to be the top exterior power of the holomorphic cotangent bundle, i.e. the space of holomorphic top forms on X.*

Accepting that the cotangent bundle is a holomorphic vector bundle, a fact that is proven much the same as proving the the ordinary tangent bundle is really a vector bundle, it is clear that K_X is a holomorphic line bundle. A section s of K_X will in local holomorphic coordinates $z = (z^i)$ look like

$$s = f(z)dz_1 \wedge ... \wedge dz_n \tag{2.28}$$

and thus the fibers K_X has dimension 1. This bundle is special in that it always exists but also because of the following proposition.

Proposition 5. *There is a one-to-one correspondence between metrics on the canonical line bundle K_X on a complex manifold X and volume forms on X compatible with the natural orientation.*

Proof. Let $V \in \bigwedge^n T^*X$ be a volume form on X compatible with the natural orientation, i.e, a real non-zero form of top dimension which in local coordinates z^i can be written

$$V = v(z)\ \frac{i}{2}dz^1 \wedge d\bar{z}^1 \wedge ... \wedge \frac{i}{2}dz^n \wedge d\bar{z}^n \tag{2.29}$$

where v is a smooth real positive function. Given a section $s = f(z)dz^1 \wedge ... \wedge dz^n$ of K_X we define the metric associated to the volume form V via

$$||s||^2 := (i/2)^n s \wedge \bar{s}/V := |f(z)|^2/v(z). \tag{2.30}$$

By the formal division in the middle expression this is clearly independent of choice of coordinates.

If we on the other hand are given a metric $||\cdot||$ then in coordinates z^i we can define at least locally a positive real function $v(z) = ||dz^1 \wedge ... \wedge dz^n||$. Under holomorphic coordinate changes, this function transforms precisely in such a way that $V := v(z)\frac{i}{2}dz^1 \wedge d\bar{z}^1 \wedge ... \wedge \frac{i}{2}dz^n \wedge d\bar{z}^n$ becomes a global volume form. Additionally performing these constructions twice yields back the same volume form or metric on K_X. Thus we get the one-to-one correspondence.

□

To recap we can divide compact complex manifolds into four kinds depending on which class in Proposition 4 the canonical bundle is a member of. In fact this division is important in the study of complex projective varieties. When the canonical bundle is positive the variety is called canonically polarized, when the canonical bundle is negative the variety is called Fano, when it is is trivial it is called Calabi-Yau. In the last case there is in the authors knowledge no name but they can be considered general varieties and the other cases rather special cases. As we have stated before in this thesis we are only concerned with the Fano case.

2.4 Kähler–Einstein metrics

There are several reasons to consider Kähler–Einstein metrics. One of the motivations is to look for a canonical metric on a complex manifold. We ask if there is some natural restrictions we can demand on a metric so that we end up with a unique one. In that case the metric could contain plenty of information about the underlying manifold.

It is natural to at least restrict to Kähler metrics, remembering that they can be thought of as Riemannian metrics with a natural compatibility with the complex structure. We start with a Kähler metric, represented by the Kähler form ω on a complex manifold X. The Kähler form induces a natural volume form on X given by $\omega^n/n!$. This gives a natural metric on K_X by Proposition 5. We can then consider the curvature form of this metric. We will denote the negative of this curvature form by Ric ω and in local coordinates we get

$$\text{Ric } \omega = -\frac{i}{2\pi}\partial\bar{\partial}\log\frac{\omega^n/n!}{(i/2)^n dz \wedge d\bar{z}}. \tag{2.31}$$

In fact, this is essentially the Ricci tensor from Riemannian geometry [8, Section. 8.5.2]. It is remarkable that in Kähler geometry, the Ricci curvature takes this simple form. Since the Ricci form is itself a J-invariant symplectic form an arguably natural thing to impose now is

$$\text{Ric } \omega = \lambda\omega \tag{2.32}$$

where after normalization we have $\lambda \in \{-1, 0, 1\}$. Such metrics will be called Kähler–Einstein metrics. Fixing the underlying complex manifold, we cannot choose λ. In fact, if $\lambda = 0$ then the first Chern class of K_X is 0, something which is equivalent to K_X

being trivial, but we wont show it here. If instead $\lambda = -1$, then K_X is positive in the sense that the metric induced from the volume form $\omega^n/n!$ has positive curvature form ω. If $\lambda = 1$, then K_X^* is positive for the same reason. Comparing with Proposition 4 this covers the different cases for K_X except the last one, where no solution at all can exist for (2.32). In this case one looks for other metrics instead, such as constant scalar curvature metrics. Note that metrics solving (2.32) also have constant scalar curvature. For $\lambda \in -1, 0$ there are theorems guaranteeing the unique existence of such a metric. The case $\lambda = 1$, when K_X^* is positive, is more difficult and a fairly recent theorem settled that solutions exists under a certain 'stability' condition [4]. Here the solutions are unique up to automorphisms connected to the identity. We will get back to existence in the chapter on the variational formulation.

We return to the Kähler–Einstein equation with $\lambda = 1$,

$$\text{Ric}\ \omega = \omega. \tag{2.33}$$

One way to think about this equation is as a certain relation to hold not for a Kähler metric, but for a metric on the line bundle K_X^*. Recall how $\omega^n/n!$ is a metric on K_X. The weight function ϕ of the dual metric on K_X^* in a trivialization from holomorphic coordinates z_i is given by

$$\phi = -\log \frac{\omega^n/n!}{(i/2)^n dz \wedge d\bar{z}}. \tag{2.34}$$

To end up with an equation for ϕ instead of ω we can express the metric on the line bundle in terms of ϕ in two ways. Computing the length of the trivializing section we get first, by definition of ϕ

$$||dz||^2 = e^{-\phi}. \tag{2.35}$$

But we also defined $|| \cdot ||$ to be dual to the metric associated to the volume form $\omega^n/n!$, i.e,

$$||dz||^2 = \frac{\omega^n/n!}{(i/2)^n dz \wedge d\bar{z}}. \tag{2.36}$$

By the Kähler–Einstein equation, ω is just the curvature form of $||\cdot||$, thus we have

$$||dz||^2 = \frac{(i\partial\bar{\partial}\phi)^n/n!}{(2\pi)^n(i/2)^2 dz \wedge d\bar{z}}. \tag{2.37}$$

Putting it all together the partial differential equation for ϕ reads

$$(\partial\bar{\partial}\phi)^n = \pi^n n! e^{-\phi} dz \wedge d\bar{z}. \tag{2.38}$$

We can further massage the left hand side into

$$\begin{aligned}
(\partial\bar{\partial}\phi)^n &= \left(\sum_{i,j=1}^{n} \frac{\partial^2\phi}{\partial z_i \partial \bar{z}_j} dz_i \wedge d\bar{z}_j\right)^n & (2.39)\\
&= \sum_{\sigma,\tilde{\sigma}\in S_n} \prod_{i=1}^{n} \frac{\partial^2\phi}{\partial z_{\sigma(i)} \partial \bar{z}_{\tilde{\sigma}(i)}} dz_{\sigma(i)} \wedge d\bar{z}_{\tilde{\sigma}(i)} & (2.40)\\
&= \sum_{\sigma,\tilde{\sigma}\in S_n} \prod_{i=1}^{n} \frac{\partial^2\phi}{\partial z_i \partial \bar{z}_{\sigma^{-1}\tilde{\sigma}(i)}} dz_i \wedge d\bar{z}_{\sigma^{-1}\tilde{\sigma}(i)} & (2.41)\\
&= n! \sum_{\sigma\in S_n} \prod_{i=1}^{n} \frac{\partial^2\phi}{\partial z_i \partial \bar{z}_{\sigma(i)}} dz_i \wedge d\bar{z}_{\sigma(i)} & (2.42)\\
&= n! \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} \frac{\partial^2\phi}{\partial z_i \partial \bar{z}_{\sigma(i)}} dz_1 \wedge d\bar{z}_1 \wedge ... \wedge dz_n \wedge d\bar{z}_n & (2.43)\\
&= n! \det\left(\frac{\partial^2\phi}{\partial z_i \partial \bar{z}_j}\right)_{i,j} dz \wedge d\bar{z} & (2.44)
\end{aligned}$$

where in step 2, S_n denotes the permutation group of n elements. In step 3, the terms are reordered and no signs are present since swapping pairs of one-forms yields no signs. In step 4, it is used that after fixing a permutation σ', there are exactly $n!$ choices of pairs of permutations σ and $\tilde{\sigma}$ such that $\sigma' = \sigma^{-1}\tilde{\sigma}$. The permutations are then renamed. In step 5, a sign appears since swapping only the $d\bar{z}_j$ terms while fixing the non barred produces a sign. The last step is simply the definition of the determinant and we use the shorthand notation $dz \wedge d\bar{z}$ for our preferred volume form. Going back the Kähler–Einstein equation we get

$$\det\left(\frac{\partial^2\phi}{\partial z_i \partial \bar{z}_j}\right)_{i,j} dz \wedge d\bar{z} = C e^{-\phi} dz \wedge d\bar{z} \tag{2.45}$$

with $C = \pi^n$. In fact, ϕ satisfying the above equation for any C makes $\omega = dd^c\phi$ satisfy the Kähler–Einstein equation since different choices of C only produce an additive shift in ϕ, not affecting the Kähler form. This is the form of the KE equation we will use, in particular with C chosen so that the right hand side is a probability measure.

3

Toric complex manifolds

Complex projective space is special in that it admits natural charts with domains biholomorphic to $\mathbb{C}^{*n}$ covering the entire manifold except lower dimensional parts. This makes computations simple and as we will see, it will also make the connection between Kähler–Einstein metrics and optimal transport. The class of complex manifolds admitting a similar structure are toric varieties.

Definition 6. *A toric complex manifold X of dimension n is a complex manifold with a holomorphic action of $T_{\mathbb{C}}^n$ acting on it. Additionally, the action admits a dense, open, free orbit in X.*

Here, $T_{\mathbb{C}} = \mathbb{C}^* $ is the complexified torus $\mathbb{C}^* = \mathbb{C}/\{0\}$. We begin with the following proposition, giving us the chart structure whose existence we have alluded to.

Proposition 6. *Let X be a toric complex manifold and $\mathcal{O}$ a dense, open, free orbit of $T_{\mathbb{C}}^n$. Then $\mathcal{O}$ is biholomorphic to $T_{\mathbb{C}}^n$.*

Proof. Fix a point $p \in \mathcal{O}$. Define $F : \mathbb{C}^{*n} \to \mathcal{O}$, $g \mapsto g \cdot p$. The map is holomorphic and surjective by definition. To prove injectivness assume for $g, \tilde{g} \in \mathbb{C}^{*n}$ that $F(g) = F(\tilde{g})$. Then $g \cdot p = \tilde{g} \cdot p$, but then $\tilde{g}^{-1} \cdot g \cdot p = p$ and since $\mathbb{C}^{*n}$ acts freely on $\mathcal{O}$ we have $g = \tilde{g}$. Thus F is bijective and holomorphic and thus biholomorphic, a standard fact from multivariate complex analysis, shown in [9, p. 19]. □

Thus we can view a compact toric complex manifold as a compactification of the non-compact space $\mathbb{C}^{*n}$. We will henceforth identity $\mathcal{O}$ with $\mathbb{C}^{*n}$ using the biholomorphism writing $\mathbb{C}^{*n} \subset X$. Additionally, this biholomorphism is trivially equivariant with respect to the action so that acting with an element $\lambda \in \mathbb{C}^{*n}$ on a point in $\mathcal{O}$ is the same as action by multiplication with the corresponding object in $\mathbb{C}^{*n}$.

We are interested in solving the Kähler–Einstein equation (2.32) on toric complex manifolds, specifically with $\lambda = 1$. Thus K_X^* will be positively curved. It is a general fact of complex geometry that positive line bundles possess many holomorphic sections, while negative line bundles do not possess any sections. We will use this to our advantage and study $H^0(X, K_X^*)$[1], the space of global sections of K_X^*. From Hodge theory we know that $H^0(X, K_X^*)$ is finite dimensional [9, ch. 0]. It is a vector space, and $\mathbb{C}^{*n}$ acts linearly on it via pushforward and makes it a representation module under the group $\mathbb{C}^{*n}$. We will use some elementary representation theory here. Firstly, $\mathbb{C}^{*n}$ is abelian and therefore every irreducible representation is one-dimensional and, secondly, we have a split of the space of global section into irreducible representation modules of $\mathbb{C}^{*n}$ as

$$H^0(X, K_X^*) = \bigoplus_{k \in I} s_k \tag{3.1}$$

for k in some index set I. Over $\mathbb{C}^{*n} \subset X$, K_X^* is trivialized and thus we can describe the action of $\mathbb{C}^{*n}$ easily. For $\lambda \in \mathbb{C}^{*n}$ we must have

$$\lambda \cdot s_k = f_k(\lambda) s_k \tag{3.2}$$

for some holomorphic function f_k on $\mathbb{C}^{*n}$. This because the action should be holomorphic and linear. f_k also has to satisfy, for any two $\lambda, \lambda' \in \mathbb{C}^{*n}$

$$f_k(\lambda\lambda') = f_k(\lambda) f_k(\lambda') \tag{3.3}$$

and has to be one the form

$$f_k(\lambda) = \lambda^{p_k} := \lambda_1^{p_{k,1}} ... \lambda_n^{p_{k,n}} \tag{3.4}$$

where p_k is a point in $\mathbb{C}^n$. Since the complex logarithm is not well defined on all of $\mathbb{C}^*$ and f_k should be well-defined on $\mathbb{C}^{*n}$ we need in fact $p_k \in \mathbb{Z}^n \; \forall k$. Henceforth we will sometimes index the sections or their local defining functions with an index i and sometimes with the corresponding point p.

[1] The notation is from sheaf theory where $H^n(X, \mathcal{F})$ for a sheaf $\mathcal{F}$ is the sheaf cohomology groups and the 0-th group corresponds to the global sections.

Working over the orbit $\mathbb{C}^{*n}$ we can be quite explicit and use the biholomorphism to see that in fact any section can be written

$$s_p = g(z)\frac{\partial}{\partial z_1} \wedge ... \wedge \frac{\partial}{\partial z_n} \tag{3.5}$$

with g holomorphic on $\mathbb{C}^{*n}$. Since by definition of the pushforward action, $\lambda \cdot \frac{\partial}{\partial z_i} = \lambda^{-1}\frac{\partial}{\partial z_i}$, we can define an invariant section s_0 over $\mathbb{C}^{*n}$ by taking $g(z) = z_1 z_2 ... z_n$. Using this non-zero section to trivialize K_X^* over $\mathbb{C}^{*n}$, the sections s_p, knowing that they are of the form (3.2) with f_p of the form (3.4), can be written, after perhaps dividing by an overall constant,

$$s_p = f_p(z)s_0. \tag{3.6}$$

A bit more can be said of $H^0(X, K_X^*)$.

Proposition 7. *Assume an integer lattice point $\tilde{p}$ is a convex combination of the points $\{p_k\}_k$, i.e, there are real numbers t_k such that $\sum_k t_k = 1$ and*

$$\tilde{p} = \sum_k t_k p_k. \tag{3.7}$$

Then $s_{\tilde{p}}$ is a section of K_X^*.

Proof. Consider $|f_{\tilde{p}}(z)|$ and the Log-map on $\mathbb{C}^{*n}$, $w_i = \log(|z_i|^2)$. We get

$$\begin{aligned}
\mathrm{Log}_*|f_{\tilde{p}}|(w) &= \exp(w \cdot \tilde{p}/2) = \exp(\frac{1}{2}w \cdot \sum_i t_i p_i) && (3.8)\\
&= \prod_i \exp(p_i \cdot w/2)^{t_i} \leq \prod_i \max_j \exp(p_j \cdot w/2)^{t_i} && (3.9)\\
&= \max_j \exp(p_j \cdot w/2)^{\sum_i t_i} = \max_j \exp(p_j \cdot w/2) && (3.10)\\
&\leq \max_j \mathrm{Log}_*|f_{p_j}|(w) && (3.11)
\end{aligned}$$

By the Riemann extension theorem [9, p. 9] the section $s_{\tilde{p}}$ over $\mathbb{C}^{*n}$ has to extend to all of X.

□

Now we have a fairly good idea of how the space of sections $H^0(X, K_X^*)$ looks like. There is a basis of elements that are elements of irreducible one-dimensional modules under $\mathbb{C}^{*n}$ and are of the form (3.6) over the orbit $\mathbb{C}^{*n}$. Additionally, because of the identity theorem for holomorphic functions, the points p are unique, i.e, there are no more than one section with the behavior (3.4) for each point p. Additionally, the set of points p for each section will be of the form of the integer lattice $\mathbb{Z}^n$ intersected with a convex polytope. We will use these sections to construct a special metric on K_X^*. Constructing metrics from sections is general procedure that we now describe.

Let $L \to X$ be a line bundle and $s \in H^0(X, L)$ a global section. Let e_U be a local trivializing section over $U \subset X$ and let s_U be the locally defined holomorphic function such that $s = s_U e_U$. We define a metric $|| \cdot ||$ first locally on U with respect to this trivialization by defining the weight function ϕ of the metric to be

$$\phi_U = \log |s_U|^2. \tag{3.12}$$

This works as long as s_U is non-zero, otherwise we get a singular metric. Postponing this issue for a while we can also wonder if this is well-defined globally, i.e, if it is independent of choice of trivialization. If e_V is another local trivializing section over V with $V \cap U \neq \emptyset$, then $e_U = he_V$ on $U \cap V$ for h holomorphic and non-zero. In this trivialization $s = s_V e_V = s_U h e_V$ over $U \cap V$ and we can define

$$\phi_V = \log |s_U h|^2 = \log |s_U|^2 + \log |h|^2. \tag{3.13}$$

This is precisely how two weight functions in two different trivializations but corresponding to the same metric should differ, indeed

$$||s||^2 = e^{-\phi_U} |s_U|^2 = e^{-\phi_U} |s_V|^2 |h|^{-2} := e^{-\phi_V} |s_V|^2 \tag{3.14}$$

where ϕ_V is the weight function of the metric in the trivialization over V.

It is easy to see that we can also produce metric from two global sections $s^{(1)}$ and $s^{(2)}$ by putting, for a trivialization over U, the weight function ϕ_U of the metric over U to be

$$\phi_U = \log(|s_U^{(1)}|^2 + |s_U^{(2)}|^2) \tag{3.15}$$

since it will have the same transformation relation between trivializations. The metric will be non-singular whenever the two sections have disjoint zero-set.

In the case of a toric manifold, we can define a canonical (a priori possibly singular) metric by letting

$$\phi_U = \log \sum_p |s_{p,U}|^2 \tag{3.16}$$

where s_p are the sections corresponding to the irreducible representations of $\mathbb{C}^{*n}$ defined up to normalization in (3.1). Consider ϕ in the trivialization defined by s_0 over the orbit $\mathbb{C}^{*n}$, we then have

$$\phi = \log \sum_p |z^p|^2 \tag{3.17}$$

where z^p as before is to be interpreted as $z_1^{p_1} z_2^{p_2} ... z_n^{p_n}$. For now, assume that the resulting metric is indeed non-singular i.e, the mutual zero-set of the sections s_p is empty. This is not always the case but for toric compact complex which is also Fano, i.e. has negative canonical bundle, it is as will be evident by the end of the chapter. For now we will just assume that we work with a complex manifold that has this property. From this it is also clear that there are at least two such sections since any one section will have some non-empty zero-set, otherwise the K_X^* would be trivial contradicting its positiveness.

We introduce real coordinates that what we will refer to as Log-coordinates w by putting $w = \operatorname{Log} z := (\log |z_1|^2, ..., \log |z_n|^2)$. This coordinate change is clearly not one-to-one since the pre-image of any point is a real torus but ϕ is invariant with respect to the real torus T^n embedded in $\mathbb{C}^{*n}$. It also maps $\mathbb{C}^{*n}$ to $\mathbb{R}^n$. In these coordinates we have

$$\phi(w) = \log \sum_p \exp(w \cdot p). \tag{3.18}$$

This function is convex, something we will utilize, indeed $\forall w_1, w_2 \in \mathbb{R}^n, t \in [0, 1]$

$$\phi(w_1 t + (1-t)w_2) = \log \sum_p \exp((w_1 t + w_2(1-t)) \cdot p) \tag{3.19}$$
$$= \log \sum_p \exp(w_1 \cdot p)^t \exp(w_2 \cdot p)^{1-t} \tag{3.20}$$
$$\leq \log \left(\sum_p \exp(w_1 \cdot p) \right)^t \left(\sum_p \exp(w_2 \cdot p) \right)^{1-t} \tag{3.21}$$
$$= t \log \sum_p \exp(w_1 \cdot p) + (1-t) \log \sum_p \exp(w_2 \cdot p) \tag{3.22}$$
$$= t\phi(w_1) + (1-t)\phi(w_2) \tag{3.23}$$

by Hölder's inequality using the $L^{1/t}$ and $L^{1/(1-t)}$ norms.

The gradient of this function is

$$\nabla \phi(w) = \frac{\sum_p p \exp(p \cdot w)}{\sum_p \exp(p \cdot w)} \tag{3.24}$$

This gradient is quite special, namely

Proposition 8. *The function* $\nabla \phi = \frac{\sum_p p \exp(p \cdot w)}{\sum_p \exp(p \cdot w)}$ *satisfies* $\overline{\nabla \phi(\mathbb{R}^n)} = P$ *where* P *is the convex hull* $\operatorname{Conv}(I)$ *of the set of points* $p \in I$*.*

Proof. Denote by $\tilde{I}$ the set of points $\tilde{p} \in I$ satisfying that

$$\tilde{p} \notin \operatorname{Conv}(I \setminus \{\tilde{p}\}) \tag{3.25}$$

Pick any $\tilde{p} \in \tilde{I}$. By the Hahn-Banach separation theorem there is a hyperplane $\{x \in \mathbb{R}^n : n \cdot x = c\}$ with $n, c \in \mathbb{R}^n$ dividing $\operatorname{Conv}(I \setminus \{\tilde{p}\})$ and $\tilde{p}$. We can choose the normal n such that it points towards $\tilde{p}$ i.e, we have $n \cdot \tilde{p} > c$ and $n \cdot x < c$ for any $x \in \operatorname{Conv}(I \setminus \{\tilde{p}\})$. Thus for any $r \in \mathbb{R}$ we have

$$\nabla \phi(rn) = \frac{\sum_p p \exp(rn \cdot p)}{\sum_p \exp(rn \cdot p)} = \frac{\sum_p p \exp(r(n \cdot p - c))}{\sum_p \exp(r(n \cdot p - c))} \xrightarrow[r \to \infty]{} \tilde{p} \tag{3.26}$$

using the inequalities from separation condition. Thus the set $\tilde{I}$ lies in the closure $\overline{\nabla \phi(\mathbb{R}^n)}$. It is a general fact that for any convex ϕ on $\mathbb{R}^n$ we have that $\overline{\nabla \phi(\mathbb{R}^n)}$ is convex [10, p. 227]. Thus $\operatorname{Conv}(\tilde{I}) \subset \overline{\nabla \phi(\mathbb{R}^n)}$. But looking back at the definition of $\tilde{I}$ it is clear that $\operatorname{Conv}(\tilde{I}) =$

Conv(I). We also have $\overline{\nabla\phi(\mathbb{R}^n)} \subset \text{Conv}(I)$ since the gradient is a convex combination of the points in I we allow to go to infinity in $\mathbb{R}^n$. Thus $\text{Conv}(I) = \overline{\nabla\phi(\mathbb{R}^n)}$. □

The polytope appearing will be referred to as the fundamental polytope associated with the line bundle (X, K_X^*). We can wonder whether the special metric constructed is also positively curved. It is and it follows from the following result.

Proposition 9. *The pullback map* Log_* *provides a bijection between smooth convex functions on* $\mathbb{R}^n$ *and* T^n*-invariant smooth plurisubharmonic functions on* $\mathbb{C}^{*n}$*.*

Proof. Let $f\colon \mathbb{C}^{*n} \to \mathbb{R}$ be a plurisubharmonic function i.e, $\partial\bar{\partial}f > 0$, which is also T^n-invariant. We can locally change holomorphic coordinates from the standard z-coordinates to $w = 2\log(z)$, choosing a local branch of the complex logarithm. f being T^n invariant in z-coordinates turns into f being invariant with respect to imaginary shifts in the w-coordinates. In these coordinates

$$\text{Log}(z) = \log|z|^2 = \log|e^{w/2}|^2 = \text{Re}(w). \tag{3.27}$$

Let $\text{Re}(w) := x$ and compute

$$\partial\bar{\partial}f = \frac{\partial^2}{\partial w_i \partial \bar{w}_j} f dw^i \wedge d\bar{w}^j = \frac{1}{2^{2n}} \frac{\partial^2}{\partial x_i \partial x_j} f dw^i \wedge d\bar{w}^j \tag{3.28}$$

so that the condition that f is plurisubharmonic has turned into the condition that $f(\text{Re}(\cdot))$ has positive definite Hessian, and consequently is a convex function. Going back to the original z-coordinates we have proven that $\text{Log}_* f$ is convex if f is plurisubharmonic and T^n-invariant. Conversely if f is convex on $\mathbb{R}^n$ then $w \mapsto f(Re(w))$ is plurisubharmonic by the above computation. □

The fact that the special metric we found has a special connection to the fundamental polytope is nothing special but in fact the case of all metrics as is stated in the following theorem.

Theorem 1. *Let* X *be a compact toric complex manifold. Let* $||\cdot||$ *be any positively curved metric on* K_X^* *invariant under the real torus* $T^n \in \mathbb{C}^{*n}$*. Let* φ *be its weight function with respect to the trivializing section* s_0 *on* $\mathbb{C}^{*n}$ *which is invariant under the* $\mathbb{C}^{*n}$*-action. In Log-coordinates,* φ *is convex and the gradient satisfies*

$$\overline{\nabla_w \varphi(\mathbb{R}^n)} = P \tag{3.29}$$

where P *is the fundamental polytope associated to* (X, K_X^*)*.*

We postpone the proof a bit and present an interlude of some convex analysis.

Definition 7. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. We define the subgradient*[2] $\partial f(x_0)$ *of f at x_0 to be the set of vectors y such that*

$$f(x_0) + \langle y, x - x_0 \rangle \leq f(x) \quad \forall x. \tag{3.30}$$

We define the image $\partial f(A)$ of a set A to be the image of the set valued function ∂f.

The subgradient is a useful generalization of the ordinary gradient for differentiable convex functions but also a handy characterization of the gradient as can be seen from the following result.

Proposition 10. *If $f : \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable then $\partial f(x_0) = \{\nabla f(x_0)\}$.*

Another useful results is that

Proposition 11. *for any $f : \mathbb{R}^n \to \mathbb{R}$ convex $\overline{\nabla f(\mathbb{R}^n)}$ is a convex set.*

We refer to [10] for proofs.

Lemma 1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and Y a convex set. Then*

$$\partial f(\mathbb{R}^n) \subset Y \iff \exists C : f(x) \leq \sup_{y \in Y} y \cdot x + C. \tag{3.31}$$

Proof. Assume $\partial f(\mathbb{R}^n) \subset Y$. We have by the definition of the subgradient that for all $y \in (\partial \phi)(x_0)$ that $\forall x$

$$f(x_0) + \langle y, (x - x_0) \rangle \leq f(x). \tag{3.32}$$

Picking $x = 0$ we get

$$f(x_0) \leq \sup_{y \in \partial f(x_0)} y \cdot x_0 + f(0) \leq \sup_{y \in Y} y \cdot x_0 + f(0). \tag{3.33}$$

[2]Not to confuse with the ∂ appearing as a derivative of forms. But since that only appear for functions on a complex space and subgradients will only appear on convex functions on a real space there should be no confusion.

Now assume instead that

$$f(x) \leq \sup_{y \in Y} y \cdot x + C \tag{3.34}$$

for some constant C and all $x \in \mathbb{R}^n$. Pick a point $x_0 \in \mathbb{R}^n$ and $z \in (\partial f)(x_0)$. We have by the definition of subgradient that

$$f(x_0) + z \cdot (x - x_0) \leq f(x) \quad \forall x \tag{3.35}$$

and using the assumed inequality

$$f(x_0) + z \cdot (x - x_0) \leq \sup_{y \in Y} y \cdot x + C \quad \forall x \tag{3.36}$$

which after rearranging reads

$$0 \leq \sup_{y \in Y} (y - z) \cdot x + z \cdot x_0 - f(x_0) + C \tag{3.37}$$

which can only hold if $z \in Y$, otherwise we could use the freedom in x to make the expression arbitrarily negatively large. Since x_0 and z where arbitrary chosen we get the desired inclusion of sets. □

Now we are fit to prove Theorem 1.

Proof. Theorem 1. We know that the function $\phi(w) = \log \sum_p \exp(w \cdot p)$ is the weight function in Log-coordinates over $\mathbb{C}^{*n}$ of a global positively curved metric. It is convex and satisfies $\overline{\nabla \phi(\mathbb{R}^n)} = P$. Any other positively curved metric has a weight function φ which is convex in Log-coordinates by Proposition 9. Since $\mathbb{C}^{*n}$ is dense in X the weight function φ of any other metric has to satisfy

$$\phi = \varphi + \mathcal{O}(1), \tag{3.38}$$

otherwise the metric will have singularities. Thus we have $\varphi \leq \phi + C$ for some constant C. Trivially $\nabla \varphi(\mathbb{R}^n) \subset \overline{\nabla \varphi(\mathbb{R}^n)}$ and the latter set is convex by Proposition 11. Let for a convex set K, $h_K(x) := \sup_{y\in K}\langle x, y\rangle$. Using Lemma 1 on ϕ we have $\phi \le h_{\overline{\nabla\phi(\mathbb{R}^n)}} + C'$ for some constant C'. Thus $\varphi \le h_{\overline{\nabla\phi(\mathbb{R}^n)}} + C + C'$. Again by Lemma 1 we get $\nabla\varphi(\mathbb{R}^n) \subset \overline{\nabla\phi(\mathbb{R}^n)}$. Exchanging the role of ϕ and φ in the above arguments we conclude that the set $\overline{\nabla\phi(\mathbb{R}^n)}$ is invariant of the choice of metric and thus is always equal to the fundamental polytope P.

□

The utility of this theorem is that is gives us a condition on a convex function ϕ on $\mathbb{R}^n$ which is equivalent to it being the weight function of a global positively curved, T^n-invariant metric on K_X^*. Among all such metrics we want to single out a Kähler–Einstein metric, thus we wish to write the KE equation in the Log-coordinates. We start with the Kähler–Einstein equation (2.45) written in terms of the unknown weight function ϕ in local holomorphic coordinates $w^i = 2\log z$,

$$\det\left(\frac{\partial^2\phi}{\partial w_i \partial \bar{w}_j}\right)_{i,j} dw \wedge d\bar{w} = Ce^{-\phi} dw \wedge d\bar{w}. \tag{3.39}$$

In Proposition 9 where convexity of the weight functions were proved, we deduced that

$$\frac{\partial^2}{\partial w_i \partial \bar{w}_j}\phi = \frac{1}{2^{2n}} \frac{\partial^2}{\partial x_i \partial x_j}\phi \tag{3.40}$$

where $x_i = \mathrm{Re}(w_i)$. Thus we get

$$\det\left(\frac{\partial^2\phi}{\partial x_i \partial x_j}\right)_{i,j} = Ce^{-\phi} \tag{3.41}$$

where C was conveniently redefined. While the w^i-coordinates only existed locally due to the nature of the complex logarithm, the x^i-coordinates are global on $\mathbb{R}^n$ and solving the above PDE on all of $\mathbb{R}^n$ gives us then a Kähler–Einstein metric on $\mathbb{C}^{*n}$ given that we restrict to convex ϕ. Since we wish to get a global metric on X, a compactification of $\mathbb{C}^{*n}$, we use Theorem 1 to essentially provide boundary conditions to the above PDE. The full formulation of the problem becomes now

$$\det\left(\frac{\partial^2\phi}{\partial x_i \partial x_j}\right)_{i,j} = Ce^{-\phi} \tag{3.42}$$

$$\nabla\phi(\mathbb{R}^n) = P \tag{3.43}$$

$$\phi \in \mathcal{C}(\mathbb{R}^n) \cup C^\infty(\mathbb{R}^n) \tag{3.44}$$

where C is some real constant which can be chosen arbitrarily and $\mathcal{C}$ denotes convex functions. It is the purpose of the next chapter to introduce the area of optimal transport and relate it to this problem. But first, a natural observation to make is that a priori the above problem makes sense for plenty of different polytopes P, but for which polytopes is there a corresponding compact toric Fano variety having P as its fundamental polytope?

Thus we will present the theory in part opposite from the theory presented from above where we started with a compact toric manifold and a toric line bundle and showed the connection to a convex polytope. Here we will start with a convex polytope and construct a compact complex manifold with a relation to P similar to the above case.

Definition 8. *A polytope $P \subset \mathbb{R}^n$ is a Delzant polytope if it satisfies*

- *The vertices are integer points.*
- *At any vertex v there are precisely n different $n-1$-dimensional faces meeting at v.*
- *At any vertex v one can choose normal vectors to the adjacent faces with integral coefficients making up a basis of the lattice $\mathbb{Z}^n$.*

Given a Delzant polytope P we will construct a complex manifold. Let $\{p_1, ..., p_N\}$ be the integer lattice points in P. Consider the map

$$\iota : \mathbb{C}^{*n} \to \mathbb{C}\mathrm{P}^{N-1} \tag{3.45}$$

$$z \mapsto [z^{p_1} : ... : z^{p_N}] \tag{3.46}$$

embedding the complex torus into complex projective space. For an introduction to complex projective space, see chapter 6. We define a compact complex toric manifold X_P by taking the closure of the image $\iota(\mathbb{C}^{*n})$. We have to show the resulting space is smooth. We will do this by constructing charts indexed by the vertices in P. Fix a vertex v in P. By assumption we are given a basis $\{\tilde{e}_i\}$ of $\mathbb{Z}^n$ of normal vectors to the faces adjacent to v. Consequently we can find a linear transformation A with integral matrix entries such that

$Ae_i = \tilde{e}_i \; \forall i$. Thus we have $P = v + AP'$ where P' is a polytope that coincides with the positive quadrant in $\mathbb{R}^n$ close to 0. This corresponds to a coordinates change $p = v + Ap'$. Now we would like to choose coordinates $\{z'_i\}_i$ on $\mathbb{C}^{*n}$ that implements this coordinates change. To this end, consider

$$z'_i = \prod_j z_i^{A_{ji}}. \tag{3.47}$$

This coordinate change implies

$$z^p = z^{v+Ap'} = z^v z'^{p'}. \tag{3.48}$$

Denoting the coordinate change by Φ_v, we can consider $\Phi_{v*}F$ which has the form

$$\Phi_{v*}\iota(z') = [z^v : z^v z'_1 : z^v z'_2 : \ldots : z^v z'_n : z^v z'^{p'_{n+2}} : \ldots : z^v z'^{p'_N}] \tag{3.49}$$

$$= [1 : z'_1 : z'_2 : \ldots : z'_n : z'^{p'_{n+2}} : \ldots : z'^{p'_N}]. \tag{3.50}$$

The domain can be extended also to all of $\mathbb{C}^n$ and is a chart on X_P.

The next problem is to show that these charts cover X_P. To that end we introduce the moment map, m,

$$m(z) = \frac{\sum_i^N p_i |z^2|^{p_i}}{\sum_i^N |z^2|^{p_i}} \tag{3.51}$$

a priori defined on $\mathbb{C}^{*n}$. In Log-coordinates w on $\mathbb{C}^{*n}$ the moment map becomes

$$m(z) = \frac{\sum_i^N p_i |z^2|^{p_i}}{\sum_i^N |z^2|^{p_i}} \tag{3.52}$$

We have already shown in Proposition 8 using Log-coordinates that $m(w)$ maps $\mathbb{R}^n$ to the interior of P and thus $m(z)$ maps $\mathbb{C}^{*n}$ to the interior of P. Pulling back m with Φ_v we have

$$\Phi_{v*}m(z') = \frac{\sum_i^N p'_i |z'^2|^{p'_i}}{\sum_i^N |z'^2|^{p'_i}} \tag{3.53}$$

and since $p'_1 = 0$ the moment map m can be extended in a bijective and smooth fashion to all of $\mathbb{C}^n$ and its image then becomes the interior P' union with the interior of all facets of all dimensions containing v in their closure.

Now take $x \in X_P$. Then by definition there is a sequence $z^{(i)} \in \mathbb{C}^{*n}$ such that $\iota(z^{(i)}) \to x$. Consider $\{m(z^{(i)})\}$. If it converges to a point in the interior of P then we are done. Assume otherwise, then it converges to a point on the boundary of P. Pick a vertex v for which $\lim_{i\to\infty} m(z^{(i)})$ lies in the interior of an adjacent facet to v (this facet does not need to be $(n-1)$-dimensional). By the previous observation, using the moment map, the sequence $z'^{(i)} = \Phi(z^{(i)})$ remains bounded and thus by compactness, after perhaps passing to a subsequence, converges in $\mathbb{C}$ to some point $z'^{(\infty)}$. By continuity $x = \Phi_{v*}\iota(z'^{(\infty)})$ so that x indeed lies in the image of some chart. In fact, we see that the range[3] U_v of the chart F_v can be chosen to be the union of, the vertex v, all adjacent faces of every dimension, and the " "interior" $\mathbb{C}^{*n}$.

X_P has a natural toric line bundle L which is given by pulling back the hyperplane bundle $\mathcal{O}(1)$ (see chapter 6) on $\mathbb{C}\mathrm{P}^{N-1}$ to X_P via the embedding ι, i.e, $L = \iota^*(\mathcal{O}(1))$. Since sections of $\mathcal{O}(1)$ are the one-homogeneous functions on $\mathbb{C}\mathrm{P}^{N-1}$ we can take a basis of $H^0(\mathbb{C}\mathrm{P}^{N-1}, \mathcal{O}(1))$ to be the set of homogeneous coordinate functions. Thus $\iota^*(\mathcal{O}(1))$ is trivial over $\mathbb{C}^{*n} \hookrightarrow X_P$ and $H^0(X_P, \iota^*(\mathcal{O}(1)))$ is represented over the trivialization precisely by the monomials $\{z^{p_i}\}_{i=1}^N$.

Let us now consider first a compact toric Fano complex manifold with associated polytope P as in the beginning of the chapter. If the polytope P is Delzant we could construct the above described toric complex manifold X_P. A natural question to ask is if X and X_P has to be biholomorphic. Note that over $\mathbb{C}^{*n}$ embedded in both X and X_P, the sections of K_X^* and L are represented by the same holomorphic functions. Consequently using the sections of K_X^* on X we can embed X into $\mathbb{C}\mathrm{P}^{N-1}$ in the exact same way as X_P and consequently they are indeed biholomorphic. We also left a detail when constructing the metric ϕ in (3.16). We needed that the sections had empty mutual zero-set. Knowing that X is really the same as X_P this is now simple to see from the above arguments when we proved that the embedding F extends to all of X_P. A last detail that we will not prove is that if P is not Delzant, then X_P will not be smooth, thus if we start with X smooth, the associated polytope will be Delzant. Thus all smooth toric Fano complex manifolds as we define them are of the form X_P, or to be more general, really any smooth complex manifold equipped with a positive line-bundle are of the form X_P.

A next natural question ask is what more we now about the polytope P if we want L on

[3]One often talks about the domain of a coordinate chart but here we are considering charts to be defined from subsets of $\mathbb{C}^n$ to the complex manifold in question and not the other way around, thus we talk about coordinates ranges instead.

X_P to be the anti-canonical bundle so that X_P becomes Fano. The first observation, is that the section s_0 over $\mathbb{C}^{*n} \hookrightarrow X$ is in fact global.

Proposition 12. *Let X be a toric complex compact manifold. Then there is a unique, global, torus-invariant section s_0 of the anti-canonical bundle K_X^*, up to multiplication by $\mathbb{C}^*$.*

Proof. Pick a holomorphic vector $v \in T_1\mathbb{C}^{*n}$ where e is the identity in $\mathbb{C}^{*n}$. Denote the action of $\mathbb{C}^{*n}$ on X by $G : \mathbb{C}^{*n} \times X \to X$; $(\lambda, x) \mapsto \lambda \cdot x$. We define a holomorphic vector field V via

$$V_x = \mathrm{d}[G(\cdot, x)](v) \tag{3.54}$$

for any $x \in X$. In other words, we differentiate the action with respect to the torus variable and use the differential to pushforward a holomorphic vector from the tangent space[4] at the unit in $\mathbb{C}^{*n}$ to the different points in X. Since the action is holomorphic the vector field must be as well. Indeed because v was holomorphic and G is holomorphic in the first variable, V will be a holomorphic vector at each point, but G is also holomorphic in the second component and thus V also varies over X in a holomorphic fashion. To see the invariance under the $\mathbb{C}^{*n}$-action lifted to TX compute for $\lambda \in \mathbb{C}^{*n}$ and $x \in X$

$$\lambda \cdot V_x = \mathrm{d}G(\lambda, \cdot)(V_x) = \mathrm{d}G(\lambda, \cdot)[\mathrm{d}G(\cdot, x)(v)] \tag{3.55}$$

$$= \mathrm{d}G(\lambda, G(\cdot, x))(v) = dG(\cdot, G(\lambda, x))(v) = V_{\lambda \cdot x}. \tag{3.56}$$

Choosing n independent vectors $v_i \in T_1\mathbb{C}^{*n}$ we can construct n vector fields $V_1, ..., V_n$ such that at least at some point they are independent, just pick a point in the dense, transitive orbit. Consequently

$$s_0 = V_1 \wedge ... \wedge V_n \tag{3.57}$$

is a holomorphic, non-trivial, toric-invariant, global section of K_X^*.

□

[4] in fact the Lie algebra of $\mathbb{C}^{*n}$

Theorem 2. *Let X_P be a compact toric Fano manifold. Then the fundamental polytope is given by inequalities of the form*

$$l_F \cdot p \geq -1. \tag{3.58}$$

where l_F is an inward pointing normal vector defining each face F of P which is primitive in the sense that the greatest common divisor of all components is 1.

Proof. Begin with the following observation. The section s^v for a vertex v is non-vanishing in the coordinates chart ι_v. Indeed consider the sections of $\iota_v^*(\mathcal{O}(1))$. Using that the sections of $\mathcal{O}(1)$ are given by the homogeneous coordinate functions we get

$$H^0(\mathbb{C}^n, \iota_v^*(\mathcal{O}(1))) = \{1, z_1', z_2', ..., \}. \tag{3.59}$$

Writing the sections instead in the chart over $\mathbb{C}^{*n}$ but in the same order we get

$$H^0(\mathbb{C}^{*n}, \iota^*(\mathcal{O}(1))) = \{z^v, ...\} \tag{3.60}$$

thus s_v is represented by simply 1 over $\mathbb{C}^n$ using ι_v and thus non-vanishing on the divisors corresponding to the neighboring facet, which in these coordinates are simply $z_i' = 0$ $i = 1, ..., n$. Now consider another section s_p. We have

$$s_p = z^p s_0 = z^v z'^{p'} s_0 = z'^{p'} s_v. \tag{3.61}$$

Thus for a divisor Δ_F for a face F given by $z_i' = 0$ in the ι_v coordinates, the order of vanishing (defined with by local data, easily proven to be coordinate invariant) $\mathrm{ord}_{\Delta_F}(s_p)$ of s_p is the i:th component of p'. If we introduce the normal vector l_F' of the face F' of P' in the primed coordinates then we can write this as

$$\mathrm{ord}_{\Delta_F}(s_p) = l_F' \cdot p'. \tag{3.62}$$

But if $p = v + Ap'$ then l_F is related to l_F' by $l_F = l_F' A^{-1}$ so that

$$l_F' \cdot p' = l_F A \cdot p' = l_F \cdot (p - v) = l_F \cdot p + l_F \cdot (-v). \tag{3.63}$$

Next note that in (3.61) , by choosing $p = 0$, that $l_F \cdot (-v)$ is the order of vanishing of s_0. Thus

$$\mathrm{ord}_{\Delta_F}(s_p) = l_F \cdot p + \mathrm{ord}_{\Delta_F}(s_0). \tag{3.64}$$

Next we compute the order of vanishing of s_0 in another way. We know that s_0 is torus invariant. Lets see what that means in the primed coordinates z_i' around some vertex v. The torus action in primed coordinates is

$$\lambda \cdot z_i' = \prod_j (\lambda \cdot z_i)^{A_{ij}} = \prod_j (\lambda_i z_i)^{A_{ij}} = \prod_j \lambda_i^{A_{ij}} \prod_j z_i^{A_{ij}}. \tag{3.65}$$

Thus since A_{ij} is invertible if s_0 is to be invariant we need

$$s_0 = C z_1' ... z_n' \frac{\partial}{\partial z_1'} \wedge ... \wedge \frac{\partial}{\partial z_n'} \tag{3.66}$$

where $C \in \mathbb{C}$ is some constant. Thus s_0 clearly has order of vanishing 1 for all faces.

Next use that all global sections s_p are represented by precisely the integer lattice point in P, and thus we need for all $p \in P \cap \mathbb{Z}^n$, that

$$\mathrm{ord}_{\Delta_F}(s_p) \geq 0 \tag{3.67}$$

but using what we have proved above this is simply

$$l_F \cdot p \geq -1 \tag{3.68}$$

and this holds for every face F. □

There is a converse to the previous theorem that we will only sketch a proof of. In fact for any Delzant P of the form as in Theorem 2, X_P will be Fano with P as its fundamental polytope, meaning that $\iota^*(\mathcal{O}(1))$ really is the anti-canonical bundle on X_P. The idea of the proof is to use that we still have an invariant section s_0 due to the form of the inequalities. Also due to the form of the inequalities it has the form

$$s_0 = C z_1' ... z_n' s_v \tag{3.69}$$

around any vertex v, i.e, it vanishes of order one on the divisors corresponding to the faces of P. But then one can infer how sections of $\iota^*(\mathcal{O}(1))$ will transform between charts and one finds that they transform precisely like sections of $K^*_{X_P}$ and thus the bundles are isomorphic and we get the result.

4

Optimal transport

Imagine the problem of moving a distribution of mass from one configuration to another. There are generically several ways to do this. Thus we can talk about the minimization problem of finding the optimal way to perform the moving while minimizing the total cost of the procedure. We will restrict to the cost given by the squared Euclidean distance of the move performed, weighted by the mass. One way of of formalizing this problem in a fair bit of generality gives the Monge-Problem.

Definition 9 (The Monge-Problem). *Let μ and ν be probability measures on $\mathbb{R}^n$ and $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ a cost function. Consider the following minimization problem.*

$$\inf_T \int c(x, T(x))\mu \tag{4.1}$$

$$s.t.\ T : \mathbb{R}^n \to \mathbb{R}^n \text{ is a measurable map such that } T_*\mu = \nu \tag{4.2}$$

where $T_*\mu$ is the pushforward of μ under T. i.e, $T_*\mu(U) = \mu(T^{-1}U)$ for any measurable set U. Any map T satisfying the constraint will be referred to as a transport map.

The problem (4.2) is both non-linear in T and with non-linear constraints. Additionally, it can be show that the constraint set can be non-compact. One can also argue that it is quite unnatural since mass cannot be split, although it can be merged. One can consider a relaxation of the problem by allowing the splitting of mass, such a problem is given in the Kantorovich problem.

Definition 10 (The Kantorovich problem). *Let μ and ν be probability measures on $\mathbb{R}^n$. Consider the following minimization problem.*

$$C(\gamma) = \inf_{\gamma} \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y)\gamma \tag{4.3}$$

$$s.t\ \gamma \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)\ and\ (\mathrm{pr}_1)_*\gamma = \mu, (\mathrm{pr}_2)_*\gamma = \nu \tag{4.4}$$

where pr_i *is the natural projection onto the i:th component. We will denote the constraint set by* $\Pi(\gamma, \nu)$*. Any measure* γ *satisfying the set will be referred to as a transport plan.*

A natural cost function for these problems is the squared euclidean cost, $c(x,y) = ||x-y||^2$. For this cost function the theory of optimal transport is quite developed and it is sort of the cost function we will be working with. But observe that $||x-y|| = ||x||^2 - 2\langle x, y\rangle + ||y||^2$ and thus

$$\int ||x - T(x)||^2 \mu(x) \tag{4.5}$$

$$= \int ||x||^2\mu - 2\int \langle x, T(x)\rangle\mu + \int ||T(x)||^2\mu \tag{4.6}$$

$$= \int ||x||^2\mu - 2\int \langle x, T(x)\rangle\mu + \int ||x||^2 T_*\mu \tag{4.7}$$

$$= \mathrm{Var}(\mu) - 2\int \langle x, T(x)\rangle\mu + \mathrm{Var}(\nu) \tag{4.8}$$

for measures satisfying $\nu = T_*\mu$, where in the second step the measure theoretic change of variable formula and the constraint on T were used. The above computation ensures that the cost is finite for finite-variance measures by applying Cauchy-Schwarz on the middle term. But more so, since the variances are independent of the transport map we get the same optimal transport map, if we exchange the Euclidean cost to $c(x,y) = -\langle x, y\rangle$ albeit with a different optimal cost. This version of the optimal transport problem is the one we will continue with as it will be intimately connected to Kähler–Einstein metrics, but having in mind that it is in terms of the solution equivalent to squared Euclidean cost whenever both are finite can be useful for analogies. There is however a notable difference when changing from the Euclidean squared cost to $c(x,y) = -\langle x, y\rangle$. In the case where one measure has compact support it is enough that the other measure has finite average absolute deviation.

Definition 11. *For a measure* μ *on* $\mathbb{R}^n$ *its average absolute deviation* a_μ *is given by*

$$a_\mu = \int ||x||\mu. \tag{4.9}$$

Proposition 13. *Consider the cost-functional in the Monge problem for the cost function*

$c(x, y) = -\langle x, y\rangle$ for transport maps between a measure μ with finite average absolute deviation and a second measure ν with compact support. It is bounded both from above and below.

Proof. Estimate

$$\left|\int -\langle x, T(x)\rangle \mu\right| \leq \int ||x||\ ||T(x)||\mu \leq \int ||x||\mu\ ||T||_{\infty,\mu} = \int ||x||\mu\ ||x \mapsto x||_{\infty,\nu} \quad (4.10)$$

where $x \mapsto x$ is the identity map on $\mathbb{R}^n$. Since μ has absolute average deviation and ν compact support the right hand side is finite[1]. □

Proposition 14. *Consider the cost-functional in the Kantorovich problem for the cost-function $c(x, y) = -\langle x, y\rangle$. Let μ have finite average absolute deviation and ν have compact support. Then the cost-functional is bounded from above and below.*

Proof. Following closely the proof of 13 estimate for $\gamma \in \Pi(\mu, \nu)$

$$|\int -\langle x, y\rangle\gamma| \leq \int ||x||\ ||T(x)||\gamma \leq \int ||x||\gamma\ ||T||_{\infty,\gamma} \quad (4.11)$$

$$= \int ||x||\mu||T||_{\infty,\mu} = \int ||x||\mu||x \mapsto x||_{\infty,\nu} \quad (4.12)$$

□

With these bounds, existence of a solution to the Kantorovich problem with cost function $c(x, y) = \langle x, y\rangle$ can be proved. The proof follows a structure which is common for variational problems and we will encounter it again later for a variational problem related to the Kähler–Einstein equation.

Theorem 3. *Let μ be a measure with finite average absolute deviation and ν a measure with compact support and let $c(x, y) = -\langle x, y\rangle$. Then there exist a probability measure γ attaining the infimum in the Kantorovich problem.*

Proof. We will prove existence by assuming a sequence of transport plans approaching the minimum, and then show that this sequence approaches a minimizing transport plan.

[1]When we write $||\cdot||_p$ for $p \in [1, \infty]$ we mean the L_p norm on functions on $\mathbb{R}^n$ on $\mathbb{R}$ while with $||\cdot||$ we always mean the finite-dimensional euclidean norm.

We equip the space of probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ with the weak topology. i.e, for a sequence of probability measure $\{\gamma_k\}_k$ $\gamma_k \to \gamma$ if

$$\forall f \in \mathcal{C}_b(\mathbb{R}^n \times \mathbb{R}^n) \int f\gamma_k \to \int f\gamma. \tag{4.13}$$

Where $\mathcal{C}_b(\mathbb{R}^n \times \mathbb{R}^n)$ are the bounded continuous functions.

Step 1: Compactness

We need to ensure that the constraint set is compact somehow. The space of probability measures on a compact subset of $\mathbb{R}^n$ is compact in the weak topology but since μ is not assumed to have compact support we are not in that case. The sequence of delta-measures $\{\delta_n\}$ with masses at consecutive natural numbers is an example of a sequence of probability measures which is easily seen to not have a weakly convergent subsequence in the space of probability measures. In this case we are saved by Prokhorov's theorem which states that if a sequence of probability measures $\{\gamma_k\}_k$ is *tight* in the sense that for every $\epsilon > 0$ there exist a compact set $K_\epsilon \subset \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n \times \mathbb{R}^n \setminus K_\epsilon} \gamma_k < \epsilon \ \forall k, \tag{4.14}$$

then there is a weakly convergent subsequence $\{\gamma_{k_l}\}$ converging to some probability measure γ [11, Theorem 5.1].

Let $\{\gamma_k\}_k$ be any sequence of transport plans. Take $\epsilon > 0$. Choose $K_\epsilon = \tilde{K}_\epsilon \times K$, K being the support of ν and with $\tilde{K}_\epsilon \subset \mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n \setminus \tilde{K}_\epsilon} \mu < \epsilon \tag{4.15}$$

This is possible since μ is a probability measure. Estimate

$$\int_{\mathbb{R}^n \times \mathbb{R}^n \setminus K_\epsilon} \gamma_k = \int_{(\mathbb{R}^n \setminus K_\epsilon) \times K} \gamma_k = \int_{(\mathbb{R}^n \setminus \tilde{K}_\epsilon)} \mu < \epsilon. \tag{4.16}$$

So any sequence of transport plans is tight and therefore posseses a weak limit which is a probability measure. Also note that by the definition of the weak topology that the map taking a probability measure on $\mathbb{R}^n \times \mathbb{R}^n$ to the pushforward measure under the projection to one of the copies of $\mathbb{R}^n$ is continuous with respect to the weak topology so that the limiting probability measure will also be a transport plan.

Step 2. Lower semi-continuity.

Next we want to show that the objective functional is lower semi-continuous to ensure that the cost of the limiting transport plans really minimizes to cost. To that end let $\{\gamma_k\}_k$ once again be a sequence of transport plans this time converging to a transport plan γ. Our main problem is that $c(x, y)$ is not bounded in our case.

Denote by K the support of ν. By assumption K is compact and thus lies in a ball of radius R for some $R > 0$. Thus we have

$$-\langle x, y\rangle \geq -||x||\ ||y|| \geq -R||x|| \quad \forall (x, y) \in \mathbb{R}^n \times K. \tag{4.17}$$

The function $x \mapsto -R||x||$ is $L^1(\gamma_k)$ $\forall k$ and $L^1(\gamma)$ since

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} R||x||\gamma_k = R \int_{\mathbb{R}^n} ||x||\mu < \infty. \tag{4.18}$$

and similarly for integration against γ.

Consider changing the cost function to $\tilde{c}(x, y) = -\langle x, y\rangle + R||x||$ instead. Since

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \tilde{c}(x, y)\gamma_k = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y)\gamma_k + R \int_{\mathbb{R}^n} ||x||\mu \tag{4.19}$$

and similarly for γ the minimization problem is not altered, only the value of the objective functional is changed by a finite amount independent of the transport plan. $\tilde{c}$ is non-negative and continuous, thus we can write it as a non-decreasing limit of continuous bounded functions $\{c_j\}_j$, then by the monotone convergence theorem

$$\int \tilde{c}\gamma = \lim_{j\to\infty} \int \tilde{c}_j\gamma = \lim_{j\to\infty} \lim_{k\to\infty} \int c_j\gamma_k = \sup_j \liminf_{k\to\infty} \int c_j\gamma_k \leq \liminf_{k\to\infty} \sup_j \int c_j\gamma_k \leq \liminf_{k\to\infty} \int \tilde{c}\gamma_k \tag{4.20}$$

where we have written the monotone limit as supremum and written the non-monotone limit as an limit inferior to be able to swap the limits with an inequality. The last step just uses that $c_j \leq c$ $\forall j$. Thus the objective functional is lower semi-continuous.

Step 3. Conclude existence

Denote by C the infimum in the Kantorovich problem. It is finite by Proposition 14. Take a sequence of transport plans γ_k such that

$$\lim_{k\to\infty} \int c\gamma_k = C. \tag{4.21}$$

By the arguments in step 1 we may assume $\gamma_k \to \gamma$ weakly with $\gamma \in \Pi(\mu, \nu)$. By step 2 we have

$$\int c\gamma \leq \lim_{k} \int c\gamma_k = C \tag{4.22}$$

and thus γ really is a minimizer. □

In the Kantorovich problem the objective function has linear dependence on γ and the constraints are linear as well. In finite dimensional linear optimization there is to any minimization problem a dual maximization where objective function and constraints change role and the solutions of the two problems are related. In this infinite-dimensional setting there is a sort of generalization.

Definition 12 (The dual Kantorovich problem)**.** *Let μ and ν be probability measures on $\mathbb{R}^n$ and $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a cost function. Consider the following maximization problem*

$$\sup_{\phi,\psi} J(\phi, \psi) = \sup_{\phi,\psi} \int \phi\mu + \int \psi\nu \tag{4.23}$$

$$s.t\ \phi, \psi \in C^\infty(\mathbb{R}^n) \text{ and } \phi(x) + \psi(y) \leq c(x, y) \tag{4.24}$$

The importance of this dual problem is the fact that in fact the optimal values of the two problems are the same.

Theorem 4. *For $c = -\langle x, y\rangle$ and μ having finite absolute deviation and ν finite support we have*

$$\sup_{\phi,\psi} J(\phi, \psi) = \inf_{\gamma} C(\gamma) \tag{4.25}$$

where ϕ, ψ and γ are in their respective constraint sets. Additionally the supremum in the dual Kantorovich problem is attained by a pair of continuous functions ϕ and ψ.

A rigorous proof of the above theorem for a large class of cost functions and assumptions on μ and ν, including our, can be found in [12, Theorem 5.9]. We will however give a sketch of a proof of the part concerned with the equality of the two problems but skipping some analysis details to at least argue why such an equality is plausible.

Sketch of proof. The idea is to try to exchange the role of the objective functional and the constraints. Begin with the Kantorovich problem

$$C = \inf_{\gamma \in \Pi(\mu,\nu)} \int c(x,y)\gamma. \tag{4.26}$$

The constraints are of equality type. The idea is to remove the constraints, and instead penalize the objective function whenever the constraints are not satisfied. We claim

$$C = \inf_{\gamma \in \mathcal{M}^+} \sup_{\phi,\psi \in \mathcal{C}(\mathbb{R}^n)} \int c(x,y)\gamma - \int \phi(x)(\mathrm{pr}_{1*}\gamma - \mu) - \int \psi(y)(\mathrm{pr}_{2*}\gamma - \nu) \tag{4.27}$$

Where $\mathcal{M}^+$ is the set of Borel measures on $\mathbb{R}^n \times \mathbb{R}^n$. Indeed if for example $\mathrm{pr}_{1*}\gamma \neq \mu$ then there is a set of measure not zero where the ϕ can be chosen such that the middle term becomes arbitrarily large and thus we must have $\mathrm{pr}_{1*}\gamma = \mu$. Similarly for the last term. We conclude that a since there are transport maps γ with finite cost they must satisfy $\gamma \in \Pi(\mu,\nu)$. This is where the non-rigorous argument enters, imagine that we can change the order of the supremum and infimum, then we get after some rewriting

$$C = \sup_{\phi,\psi \in \mathcal{C}(\mathbb{R}^n)} \inf_{\gamma \in \mathcal{M}^+} \int c(x,y) - \phi(x) - \psi(y)\gamma + \int \phi(x)\mu + \int \psi(y)\nu. \tag{4.28}$$

Consider the first term in this expression. Whenever $c(x,y) - \phi(x) - \psi(y) < 0$ on a set of measure non-zero γ can be chosen large there to make $C = \infty$. Since ϕ and ψ should also be continuous we would get

$$C = \sup_{\phi,\psi \in \mathcal{C}(\mathbb{R}^n), \phi+\psi \leq c} \int \phi(x)\mu + \int \psi(y)\nu \tag{4.29}$$

and thus the result.

□

A rigorous proof which can be applied in this special setting using other techniques like cyclical monotonicity can be found in [12, Theorem 5.9]. That theorem actually gives much more, but some of that we can show in more elementary terms. The following results characterizes the solution to the Kantorovich problem and its dual.

Theorem 5. *In the dual Kantorovich problem a solution will always be of the form ϕ, ϕ^* and in the Kantorovich problem, whenever it is equivalent to the dual problem, a minimizing transport plan γ will be supported on the graph of the subgradient of a convex function ϕ (the graph viewing the subgradient as set-valued function).*

We will use a standard easily proven fact from convex analysis.

Lemma 2. *Let ϕ be a convex function, then $x \in \partial\phi(y)$ if and only if*

$$\phi(x) + \phi^*(y) = \langle x, y\rangle. \tag{4.30}$$

Proof. Assume $y \in \partial\phi(x)$. Then by definition

$$\phi(x) + \langle y, s - x\rangle \leq \phi(s) \quad \forall s, \tag{4.31}$$

rearranging

$$\phi(x) + \langle y, s\rangle - \phi(s) \leq \langle y, x\rangle \quad \forall s \tag{4.32}$$

and by choosing $s = x$

$$\phi(x) + \phi^*(y) \leq \langle x, y\rangle. \tag{4.33}$$

But

$$\phi(x) + \phi^*(y) = \phi(x) + \sup_s \langle y, s\rangle - \phi(s) \geq \langle y, x\rangle \tag{4.34}$$

and we have the right implication. Now assume x, y is such that $\phi(x) + \phi^*(y) = \langle x, y \rangle$. Then

$$\phi(x) - \langle x, y \rangle = -\phi^*(y) \leq -\langle y, s \rangle + \phi(s) \quad \forall s \tag{4.35}$$

and thus

$$\phi(x) + \langle y, s - x \rangle \leq \phi(s) \tag{4.36}$$

completing the proof.

□

Proof. Theorem 5. By renaming $\phi,\ \psi$ to $-\phi,\ -\psi$ in the dual Kantorovich problem we have equivalently the problem

$$-C = \inf_{\phi,\psi \in \mathcal{C}(\mathbb{R}^n), \phi,\psi \geq \langle x,y \rangle} \int \phi\mu + \int \psi\nu. \tag{4.37}$$

Let (ϕ, ψ) be an admissible pair in the optimization problem. Consider the Legendre transform of ϕ,

$$\phi^*(x) = \sup_y \langle x, y \rangle - \phi(y). \tag{4.38}$$

Since

$$\phi^*(y) \geq \langle x, y \rangle - \phi(x) \quad \forall x \tag{4.39}$$

the pair (ϕ, ϕ^*) for continuous ϕ is also an admissible plan as long as ϕ^* is continuous. It is clearly lower semi-continuous but for now we do not know whether it is continuous, it could be that it takes on the value infinity somewhere. But estimate

$$\psi(y) - \phi^*(y) = \psi(y) - \left(\sup_s \langle x, s \rangle - \psi(s) \right) = \inf_s -\langle x, s \rangle + \psi(s) + \psi(y) \leq 0 \tag{4.40}$$

so ϕ^* is everywhere bounded by ψ. Thus it is continuous but also the pair (ϕ, ϕ^*) gives a lower value of the objective functional. Thus any solution to the dual Kantorovich problem, guaranteed by Theorem 4 is of the form (ϕ, ϕ^*). But more so, we can proceed in the same way again and find that in fact a solution is always of the form $(\phi^* *, \phi^*)$. Using that the Legendre transform is an involution for convex functions, we can further say that the solution will be of the form (ϕ, ϕ^*) for a continuous and convex ϕ.

For the second part of the theorem, let γ be a solution to the Kantorovich problem, by the duality theorem we have

$$\int \langle x, y\rangle \gamma = \int \phi\mu + \phi^*\nu = \int \phi(x) + \phi(y)\gamma \tag{4.41}$$

and thus $\langle x, y\rangle = \phi(x) + \phi(y)$ γ-almost everywhere. Thus the support of γ lies in the closure of the image $\overline{\partial\phi(\mathbb{R}^n)}$ of the subgradient of ϕ by Lemma 2.

□

This characterization finally lets us connect to the original Monge-Problem again whenever the measure μ is sufficiently nice.

Theorem 6. *Let as usual μ have finite absolute average deviation and μ compact support. If the measure μ additionally does not charge small sets (null sets with respect to Lebesgue measure) then $T = \nabla\phi$, defined almost everywhere, solves the Monge-Problem.*

Proof. Consider γ, the solution to the Kantorovich problem. It is supported on the graph of the subgradient of a convex function ϕ. By the Rademacher theorem from convex analysis [10, section. 25], ϕ is almost everywhere differentiable, and by Proposition 10, in these points the subgradient contains a single point which is just the ordinary gradient of ϕ. Define the set where ϕ is not differentiable by S. By assumption $\mu(S) = 0$ since S does not charge small sets and S is small. Similarly we can divide the support of γ into $\text{supp}(\gamma) = \Gamma_1 \cup \Gamma_2$ where Γ_1 is supported on the set $\overline{\partial\phi(\mathbb{R}^n \setminus S)}$ and Γ_2 is supported on $\overline{\partial\phi(S)}$. By definition of Γ_1

$$\int_{\Gamma_1} \gamma = \int_{\mathbb{R}^n \setminus S} \mu = \mu(\mathbb{R}^n) = 1. \tag{4.42}$$

Thus $\int_{\gamma_2} = 0$. Take $T = \nabla\phi$, viewed as an L^1 map. In that way T is well-defined even though $\nabla\phi$ might not be pointwisely defined. For any other transport map T', we can construct a corresponding transport plan γ', supported on the graph of T', by $\gamma' = (\mathrm{Id} \times T')_* \mu$ with $\int c(x, y)\gamma' = \int c(x, T'(x))\mu$. Thus

$$\int c(x, T(x))\mu = \int c(x, y)\gamma \leq \int c(x, y)\gamma' = \int c(x, T'(x))\mu \tag{4.43}$$

and thus T is a minimizing transport map.

□

This is the result that finally allows us to connect optimal transport to the Kähler–Einstein equation. We begin with the Monge-Problem (4.2) and assume further that the measures are absolutely continuous with respect to the Lebesgue measure, i.e, $\mu = f\mathrm{d}x$ and $\nu = gdx$. Any admissible transport map T should transport μ to ν, that is for all measurable sets U we should have

$$\int_U f(x)dx = \int_{T(U)} T_*(f\mathrm{d}x) = \int_{T(U)} g(y)dy. \tag{4.44}$$

Now we perform the variable substitution $T(x) = y$ in the last integral, skipping issues of regularity of the transport map.

$$\int_{T(U)} g(y)\mathrm{d}y = \int_U g(T(x)) \det J(T(x))dx \tag{4.45}$$

so that we have for all measurable sets U

$$\int_U f(x)\mathrm{d}x = \int_U g(T(x)) \det J(T(x))dx \tag{4.46}$$

which is satisfied if and only if

$$g(T(x)) \det[J(T(x))] = f(x) \text{ a.e.} \tag{4.47}$$

Using Theorem 6 we get

$$g(\nabla\phi(x)) \det(\frac{\partial^2\phi(x)}{\partial x_i \partial x_j})_{i,j} = f(x). \tag{4.48}$$

This is the PDE formulation of optimal transport. This PDE is already quite similar to the version of Kähler–Einstein equation (2.45) we derived for toric varieties. We continue by specifying the target measure ν to be the uniform measure on a closed convex polytope P. We also a priori state that μ is supported on the entirety of $\mathbb{R}^n$, something which will be clear eventually. This forces $\overline{\nabla\phi(\mathbb{R}^n)} = P$ and $g = 1$. Finally we put $\mu = e^{-\phi}$, which also makes it clear that μ will have support in the whole of $\mathbb{R}^n$. The resulting transport problem should satisfy the PDE

$$1/\mathrm{Vol}(P)\det\left(\frac{\partial^2\phi}{\partial x_i \partial x_j}\right)_{i,j} = e^{-\phi} \tag{4.49}$$

$$\int e^{-\phi} = 1 \tag{4.50}$$

$$\nabla\phi(\mathbb{R}^n) = P \tag{4.51}$$

This is precisely the system (3.44) we derived for torus invariant Kähler–Einstein metrics on toric varieties after choosing the constant C correctly.

Commenting on the regularity issues, we will define a solution to the transport problem related to the above PDE above to be a weak solution to the above PDE. One can then show that weak solutions of this type become real smooth solutions via the general theory of elliptic partial differential equations, but we will not be concerned with that issue here.

This derivation furnishes the connection between Kähler- -Einstein metrics on toric complex manifolds and optimal transport. We will deepen this connection by studying a certain variational problem which gives rise to the above system and which could possibly be used to prove the existence of weak solutions.

5

Variational problem

In this section we present a variational problem within the setting of optimal transport whose solutions satisfies the system (4.51) weakly. The motivation is multitude, it can be used to investigate the existence of weak solutions to the Kähler–Einstein equation on toric varieties, it can also be used to motivate a numerical algorithm to solve (4.51) where one tries to numerically solve the variational problem. Additionally, it is of theoretical interest in and of itself.

It should be noted that the existence theory of toric Kähler–Einstein metrics is already well developed. The existence theory is developed in far more general cases. In the toric case, the existence has also been solved by a variational argument in [6]. The statement of the result is that there does exists a unique Kähler–Einstein metric up to holomorphic automorphisms connected to the identity on a compact toric complex manifold if and only if the mass center of the fundamental polytope P is the origin. The variational problem here is however different from the one in [6]. There, an approach varying the Kähler potential ϕ is adopted. Our goal here is to partly develop the existence theory using a natural and clean functional with dependence on the measure $\mu = e^{-\phi}$ instead. In this work we only partially develop the theory, and the partial results presented should be considered already known with perhaps the exception of Proposition 24. The goal is also to present what a variational and probabilistic statistical mechanical approach to the solution of the weak existence theory of a non-linear partial differential equation of geometric interest could look like.

We fix, as before, a closed convex polytope $P \in \mathbb{R}^n$. We consider the uniform measure on P which we will denote in the perhaps unconventional way $1_P \mathrm{d}p$.

Definition 13. *Define the energy functional $E(\mu)$ to be the cost to optimally transport μ*

to the uniform measure on P with respect to the cost function $c(x,p) = -x \cdot p$. That is

$$E(\mu) = -C(\mu, 1_P(P)dp) = -\inf_{\gamma \in \Pi(\mu,\nu)} \int_{\mathbb{R}^n \times \mathbb{R}^n} -x \cdot p\, \gamma(x,p). \tag{5.1}$$

Viewing this energy functional as the energy of an ensemble of particles which in the large particle number limit is described by the measure μ, it is natural to introduce the entropy of the distribution μ.

Definition 14. *Define the entropy functional $D(\mu)$ to be given by*

$$D(\mu) = \int_{\mathbb{R}^n} \log(\mu/\mathrm{d}x)\mu \tag{5.2}$$

where $\mu/\mathrm{d}x$ is the Radon-Nikodym derivative or simply just density of μ with respect to the Lebesgue measure. If the density does not exist we put $D(\mu) = \infty$.

Note that this definition has a sign difference from entropy in physics. Continuing the analogy with physics, we can now construct the free energy functional $F(\mu)$ for a given inverse temperature β,

$$F(\mu) = E(\mu) + \frac{1}{\beta} D(\mu). \tag{5.3}$$

One can then consider the problem of finding the probability measure that minimizes the free energy. Avoiding for now possible issues at the boundary and existence, we should expect that the functional derivative of F should vanish at a minimum and thus we compute its derivative starting with the entropy functional.

Proposition 15. *We have for probability measures μ_0 and μ_1 that*

$$\frac{d}{dt} D((1-t)\mu_0 + t\mu_1)|_{t=0} = \int_{\mathbb{R}^n} \log(\mu_0)(\mu_1 - \mu_0) \tag{5.4}$$

and thus we say that the functional derivative $\mathrm{d}D(\mu)$ is $\log(\mu/dx)$.

Proof. Compute

$$\frac{d}{dt}D((1-t)\mu_0+t\mu_1)|_{t=0} = \frac{d}{dt}\int_{\mathbb{R}^n}\log(((1-t)\mu_0+t\mu_1)/dx)((1-t)\mu_0+t\mu_1)|_{t=0} \tag{5.5}$$
$$= \int_{\mathbb{R}^n}\frac{(\mu_1-\mu_0)/dx}{((1-t)\mu_0+t\mu_1)/dx}((1-t)\mu_0+t\mu_1)|_{t=0} \tag{5.6}$$
$$+ \int_{\mathbb{R}^n}\log(((1-t)\mu_0+t\mu_1)/dx)(\mu_1-\mu_0)|_{t=0} \tag{5.7}$$
$$= \int_{\mathbb{R}^n}(\mu_1-\mu_0)+\int_{\mathbb{R}^n}\log(\mu_0/dx)(\mu_1-\mu_0) \tag{5.8}$$
$$= \int_{\mathbb{R}^n}\log(\mu_0/dx)(\mu_1-\mu_0) \tag{5.9}$$

□

Next we consider the functional derivative of the energy functional.

Proposition 16. *The energy functional* (5.1) *is differentiable and its derivative is given by*

$$\frac{d}{dt}E(\mu+t(\nu-\mu))|_{t=0} = \int_{\mathbb{R}^n}\phi(\nu-\mu) \tag{5.10}$$

where ϕ is the Kantorovich potential appearing in the optimal transport problem of transporting μ to the uniform measure on P. Put differently the functional derivative of the energy is given by $dE(\mu)=\phi$.

Proof. Begin by noting that by Kantorovich duality

$$C(\mu) = \sup_{\phi}\int_{\mathbb{R}^n}\phi\mu+\int_P\phi^*dp. \tag{5.11}$$

We define the functional $V(\phi) = \int_{\mathbb{R}^n}\phi(\nu-\mu)$ and the family of functionals $G_t(\phi) = J(\phi)+tV(\phi)$ where $J(\phi) := J(\phi,\phi^*)$ from Definition 12. Then we have

$$\frac{d}{dt}C(\mu+t(\nu-\mu))\Big|_{t=0} = \frac{d}{dt}\Big|_{t=0}\sup_{\phi}\int_{\mathbb{R}^n}\phi(\mu+t(\nu-\mu))+\int_P\phi^*dp \tag{5.12}$$
$$= \frac{d}{dt}\Big|_{t=0}\sup_{\phi}G_t(\phi). \tag{5.13}$$

By the existence theory of optimal transport, the supremum in $g(t)$ is always attained by a unique ϕ_t.

Denoting $g(t) = \sup_{\phi} G_t(\phi)$ we compute

$$g(t) - g(0) = G_t(\phi_t) - G_0(\phi_0) \tag{5.14}$$
$$= G_0(\phi_t) + tV(\phi_t) - G_0(\phi_0) \tag{5.15}$$
$$= (G_t(\phi_0) - G(\phi_0)) + (G_0(\phi_t) - G_0(\phi_0)) + t(V(\phi_t) - V(\phi_0)) \tag{5.16}$$

For the first term we get

$$G_t(\phi_0) - G(\phi_0) = tV(\phi). \tag{5.17}$$

For the third term we get

$$V(\phi_t) - V(\phi_0) = \int_{\mathbb{R}^n} (\phi_t - \phi_0)(\nu - \mu). \tag{5.18}$$

Here note that ϕ_t lies in a compact set for t in some fixed finite interval. Thus, so does any limit point $\tilde{\phi}$ as $t \to \infty$. By continuity of $G_0(\cdot) = J(\cdot)$ we get $G_0(\tilde{\phi}) = G_0(\phi_0)$ but since ϕ is the unique maximizer of $J(\cdot)$ we get $\tilde{\phi} = \phi_0$ and thus

$$\lim_{t \to 0} V(\phi_t) - V(\phi_0) = 0 \tag{5.19}$$

by continuity of $V(\cdot)$.

For the third term notice that $G_0(\phi_t) - G_0(\phi_0) \leq 0$ since ϕ_0 is optimal. Since $g(t)$ is the supremum of affine function it is convex and thus the right and left derivatives exists [10, section. 24]. For the right derivative we get

$$\frac{d}{dt} g(t)\big|_{t=0^+} \leq V(\phi_0) \tag{5.20}$$

but by changing $t \to -t$ we get

$$\frac{d}{dt} g(t)\big|_{t=0^-} \geq V(\phi_0). \tag{5.21}$$

Again since g is convex we have $\frac{d}{dt} g(t)\big|_{t=0^-} \leq \frac{d}{dt} g(t)\big|_{t=0^-}$ and thus g is differentiable at $t = 0$ and

$$\frac{d}{dt} g(t)\big|_{t=0} = V(\phi_0) = \int_{\mathbb{R}^n} \phi_0(\nu - \mu) \tag{5.22}$$

and we conclude that in terms of functional derivatives $dE(\mu) = -dC(\mu) = -\phi$ where ϕ is the Kantorovich potential transporting μ to the uniform measure on P.

□

The two previous propositions allows us to compute the functional derivative of the free energy functional (5.3). We can conclude at least heuristically that a measure μ attaining the minimum of the free energy functional should satisfy

$$dF\big|_{\mu} = 0 \tag{5.23}$$

or using Proposition 16 and 15

$$\phi - \frac{1}{\beta}\log(\mu/dx) = 0 \tag{5.24}$$

or

$$\mu/dx = e^{-\beta\phi} \tag{5.25}$$

where ϕ is the Kantorovich potential appearing in the transport problem of optimally transporting μ to the uniform measure on P. If μ is absolutely continuous with respect to Lebesgue, which it will be if it were a finite minimizer of the free energy due to the definition of the free energy, then we know from the PDE formulation of optimal transport (4.48) that we also have

$$1/\mathrm{Vol}(P)\det\left(\frac{\partial^2\phi(x)}{\partial x_i \partial x_j}\right)_{i,j} = \mu/dx \tag{5.26}$$

so in total we have the PDE

$$1/\mathrm{Vol}(P)\det\left(\frac{\partial^2\phi(x)}{\partial x_i \partial x_j}\right)_{i,j} = e^{-\beta\phi} \tag{5.27}$$

for a convex function ϕ which also satisfies $\overline{\nabla\phi(\mathbb{R}^n)} = P$ and $\int e^{\beta\phi}\mathrm{d}x = 1$. For $\beta = 1$ we get the toric Kähler–Einstein equation (3.44) but other choices of β can be interesting as well as we will see later.

At this point one would like to conclude that the variational problem of minimizing the free energy constitute another formulation of the KE equation on toric Fano manifolds. To make such a claim rigorous one would need a couple of extra results. One would need to prove a compactness result for a space of measures μ in which one could perform the variational problem. Additionally one would need a properness result on the free energy functional to assert that a minimum exists. Additionally a convexity argument to ensure a unique minimum and finally an argument to ensure the minimum is attained at an interior point so that one really gets that the functional derivative at that point should vanish.

We begin by finding a criteria for when there is definitely no solution to the variational problem. Consider a measure μ for which the free energy is finite. Let $a \in \mathbb{R}^n$ and denote by μ_a the pushforward of μ under $x \mapsto x + a$. Compute

$$
\begin{aligned}
F(\mu_a) &= C(\mu_a) + \frac{1}{\beta} D(\mu_a) && (5.28)\\
&= -\inf_{\gamma \in \Pi(\mu_a, 1_P \mathrm{d}p)} \int -\langle x, p\rangle \gamma + \frac{1}{\beta} \int \log(\mu_a/\mathrm{d}x)\mu_a && (5.29)\\
&= -\inf_{\gamma \in \Pi(\mu, 1_P \mathrm{d}p)} \int -\langle x + a, p\rangle \gamma + \frac{1}{\beta} \int \log(\mu/\mathrm{d}x_a)\mu && (5.30)\\
&= F(\mu) + \int_P \langle a, p\rangle \mathrm{d}p && (5.31)\\
&= F(\mu) + \langle a, \bar{p}\rangle && (5.32)
\end{aligned}
$$

where $\bar{p}$ is the mass center of P. From this immediately learn two things.

Proposition 17. *The free energy functional is translation invariant if and only if the mass center of P is 0.*

Proof. Follows directly from (5.32). □

Proposition 18. *If the mass center of P is not 0 then the free energy is not bounded from below on the space of probability measures.*

Proof. Take a measure μ for which the free energy is finite. Then $F(\mu_{-n}) \to -\infty$ as $n \to \infty$. □

We conclude that the minimization problem has no solution if 0 is not the mass center of P. Later when we develop further the variational theory the last proposition will imply that the Kähler–Einstein equation on toric varieties with non-balanced associated polytope has no solution.

Comment on the case of non-balanced polytope

There is a naive remedy in the case of non-balanced polytope. We could simply consider the problem of minimizing the free energy subject to the constraint that the mass center $\bar{\mu} = 0$ to exclude the situation in Proposition 18. But there is no reason to believe that the stationary point of $F(\cdot)$ under that constraint should be the Kähler–Einstein equation. In fact this is not the case, but we can learn something by following through with this anyhow. In fact, we can use the theory of Lagrange multipliers to find the stationary equation for the problem of finding the minimum of F while fixing $\bar{\mu} = 0$. The theory of Lagrange multipliers states that we can equivalently find an extremum of

$$\tilde{F}(\mu) = F(\mu) + \langle \lambda, \int x\mu \rangle. \tag{5.33}$$

with $\lambda \in \mathbb{R}^n$ and then solve for λ using the constraint $L(\mu) = 0$ where we denote by $L(\mu)$ the functional $\mu \mapsto \int x\mu$. We fix probability measures μ and ν and the variation becomes

$$\frac{d}{dt}\Big|_{t=0} \langle \lambda, L(\mu + t(\nu - \mu)) \rangle = \langle \lambda, \int x(\nu - \mu) \rangle \tag{5.34}$$

and thus

$$d\lambda \cdot L(\mu) = \lambda \cdot x. \tag{5.35}$$

Thus putting $d\tilde{F}(\mu) = 0$ we get

$$\phi + \frac{1}{\beta} \log(\mu/\mathrm{d}x) + \langle \lambda, x \rangle = 0 \tag{5.36}$$

and thus

$$\mu = \exp(-\beta(\phi - \langle \lambda, x \rangle)). \tag{5.37}$$

Writing out $L(\mu) = 0$ we get

$$0 = \int x \exp(-\beta(\phi - \langle \lambda, x \rangle))\mathrm{d}x \tag{5.38}$$

Here we could redefine ϕ by $\tilde{\phi} = \phi - \langle \lambda, x \rangle$. Since the Monge-Ampere operator is purely second order this will not change the left hand side of the corresponding PDE and we therefore end up with precisely with our version of the Kähler–Einstein equation modulo boundary conditions. The boundary condition in terms of $\tilde{\phi}$ is

$$\overline{\nabla((\tilde{\phi} + \langle \lambda, x \rangle)(\mathbb{R}^n))} = P \tag{5.39}$$

or equivalently

$$\overline{\nabla(\tilde{\phi}(\mathbb{R}^n))} = P - \lambda. \tag{5.40}$$

If we for now assume our earlier claims that the Kähler-Einstein equation has a solution iff the mass center of P is 0 and that this happens precisely when the associated variational problem has a solution then we get here that $\lambda = \bar{p}$, the mass center of P. Going back to the geometrical setup we have found a Kähler–Einstein metric defined via $\tilde{\phi}$ on a dense subset of our toric variety X, but we already know from Theorem 1 that this metric cannot be smoothly extended to all of X and will thus have singularities. We have also deduced that the existence of these metrics follows from the existence of our ordinary Kähler–Einstein equation and thus we will go back to the variational problem.

*

To prove existence of the variational problem in the case of a balanced polytope we need to ensure that the minimization problem is well defined and thus we would like a lower bound on the free energy. We begin with one for the energy functional

Proposition 19. *If the mass center of P is* 0 *then the energy is bounded from below by 0 in the class of probability measures with finite first moment.*

Proof. Let μ be a measure with finite first moment. Estimate

$$E(\mu) - \langle \bar{\mu}, \bar{p} \rangle = - \inf_{\gamma \in \Pi(\mu, 1_P \mathrm{d}p)} \int -\langle x, p \rangle \gamma - \left\langle \int x\mu, \int_P p\mathrm{d}p \right\rangle \tag{5.41}$$

$$= \sup_{\gamma \in \Pi(\mu, 1_P dp)} \int \langle x, p \rangle \gamma - \int \langle x, p \rangle \mu \otimes 1_P \mathrm{d}p \qquad \geq 0 \tag{5.42}$$

since $\mu \otimes 1_P \mathrm{d}p \in \Pi(\mu, 1_p \mathrm{d}p)$. But since $\bar{\mu}$ is finite and $\bar{p} = 0$ we get $E(\mu) \geq 0$.

□

With Proposition 18 we have proven that the energy is bounded from below if and only if the polytope has mass center in the origin. Moving on to the entropy functional we have restricted to balanced measures for simplicity since the entropy is translation invariant anyway.

Proposition 20. *Let μ be a probability measure on $\mathbb{R}^n$ with mass center at 0 and finite average absolute deviation*

$$\int |x|\mu = a. \tag{5.43}$$

Then the following bound on the entropy holds,

$$D(\mu) \geq -n\log(a) + A_n \tag{5.44}$$

where A_n is a constant not depending on a.

The proof will consist of proving that the following multivariate generalization of the Laplace distribution minimizes the entropy under the constraint of fixed average absolute deviation.

Definition 15. *We define the multivariate Laplace distribution to be the probability measure on $\mathbb{R}^n$ with density*

$$f(x) = C_n e^{-\frac{n}{a}|x|} \tag{5.45}$$

where C_n is a normalization constant such that $\int_{\mathbb{R}^n} f dx = 1$.

We compute its normalization constant and get

$$C_n^{-1} = \int e^{-\frac{n}{a}|x|} dx = B_n \int_0^\infty r^{n-1} e^{-\frac{n}{a}|x|} \mathrm{d}r = (n-1)! \left(\frac{a}{n}\right)^n B_n \tag{5.46}$$

where we performed a hyper-spherical substitution in the integral and B_n is the constant independent of a from the angle integrals.

Computing the average absolute deviation

$$\int |x| C_n e^{-\frac{n}{a}|x|} = C_n B_n \int_0^\infty e^{-\frac{n}{a}r} r^n \mathrm{d}r = C_n B_n n! \left(\frac{a}{n}\right)^{n+1} = a \tag{5.47}$$

the entropy

$$\int \log(C_n e^{-\frac{n}{a}|x|}) C_n e^{-\frac{n}{a}|x|} \mathrm{d}x = \log(C_n) - \frac{n}{A} \int |x| C_n e^{-\frac{n}{a}|x|} \mathrm{d}x = \log(C_n) - n. \tag{5.48}$$

We are now fit to prove the bound on the entropy functional.

Proposition 20. Let μ be any probability measure on $\mathbb{R}^n$ with average absolute deviation given by a. By the definition of the entropy functional we can restrict to μ with a density g. Let f be the density of μ. Consider the functional[1]

$$D(g||f) = \int \log(g/f) g \mathrm{d}x. \tag{5.49}$$

It enjoys the property of begin positive, as is seen by

$$-D(g||f) = \int \log(f/g) g \mathrm{d}x = \mathbb{E}(\log(f/g) \leq \log(\mathbb{E}(f/g)) \tag{5.50}$$

$$= \log\left(\int_{g>0} f/g \; g \mathrm{d}x\right) \leq \log\left(\int f \mathrm{d}x\right) = 0. \tag{5.51}$$

where the expected value is with respect to the measure $f\mathrm{d}x$ but restricted to where the g is non-zero. The first inequality is the Jensen inequality using that the logarithm is concave. Using the form of the Laplace density f we get

[1]Known in information theory of the Kullback-Leibler divergence of g with respect to f. When f is the Lebesgue measure which is not a probability measure, the definition makes sense anyway and is simply the entropy functional evaluated at g.

$$0 \leq D(g||f) = \int \log(g/f) g \mathrm{d}x = \int \log(g) g \mathrm{d}x - \int \log(f) g \mathrm{d}x \tag{5.52}$$
$$= D(\mu) - \int \log(C_n e^{-\frac{n}{a}|x|}) g \mathrm{d}x \tag{5.53}$$
$$= D(\mu) - \log(C_n) \int g \mathrm{d}x + \frac{n}{a} \int |x| g \mathrm{d}x = D(\mu) - \log(C_n) + n \tag{5.54}$$

but since the entropy of the Laplace measure is simply $\log(C_n) - n$ we have showed

$$D(\mu) \geq D(f \mathrm{dx}) \tag{5.55}$$

for any probability measure μ with mass center at 0 and average absolute deviation a. Using our earlier result that $C_n^{-1} = (n-1)! \left(\frac{a}{n}\right)^n B_n$ we get

$$D(\mu) \geq -n \log(a) + n \log(n) - \log((n-1)!) - \log(B_n) \tag{5.56}$$

showing the desired dependence of the bound on a.

□

It is clear from the result that it is not enough to bound the energy form below by 0 since for a sequence of measures with increasing average absolute deviation our total bound on the free energy decreases and thus we cannot conclude that a minimizer exists in the interior and at a local minimum. Then we could no longer assert that at the minimum $\mathrm{d}F = 0$ holds and thus the connection to the Kähler–Einstein equation is lost. We thus need a better bound on the energy functional. We will be able to prove a bound of this type in dimension one, and conclude that the free energy goes to infinity whenever the average absolute deviation does so. A generalization to arbitrary dimensions is in preparation.

But still there are some results that we can do in arbitrary dimension. To guarantee a solution to the variational problem we will need a compactness argument in some topology. The topology we will use is the weak topology. i.e, we say that a sequence $\{\mu_k\}_k$ of probability measures converges weakly to μ if

$$\int_{\mathbb{R}^n} f \mu_n \to \mu \ \forall f \in \mathcal{C}_b(\mathbb{R}^n) \tag{5.57}$$

where $\mathbb{C}_b(\mathbb{R}^n)$ is the space of bounded continuous function on $\mathbb{R}^n$. If we would exchange $\mathbb{R}^n$ with a compact subset K then the resulting space of probability measures would indeed be compact as we have stated before. Recall that to get compactness in the setting of $\mathbb{R}^n$ we need to pick sequences that are *tight*, where a sequence of probability measures $\{\mu_k\}_k$ is *tight* if for every $\epsilon > 0$ we can find a compact set $K_\epsilon \subset \mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n \setminus K_\epsilon} \mu_k \leq \epsilon \ \forall k. \tag{5.58}$$

By Prokhorov's theorem, we can then extract a subsequence convergent in the weak topology in dual with $\mathcal{C}_b(\mathbb{R}^n)$ [11, Theorem 5.1].

Let us imagine that we have a bound from below on the free energy that ensures that the free energy goes to infinity whenever the average absolute deviation does so. Then in the minimization we can just as well minimize over measures with a fixed bound on their absolute average deviation. In that case we actually do have tightness.

Proposition 21. *Assume $\{\mu_k\}_k$ is a sequence of probability of measures on $\mathbb{R}^n$ with a uniform bound on their average absolute deviation. i.e,*

$$\int_{\mathbb{R}^n} |x| \mu_k < A. \tag{5.59}$$

Then the sequence is tight and consequently there is a convergent subsequence with respect to the weak topology in dual with $\mathcal{C}_b(\mathbb{R}^n)$.

Proof. Let $\epsilon \geq 0$. Let $K_\epsilon = B(0, A/\epsilon)$ the ball in $\mathbb{R}^n$ centered at the origin and with radius A/ϵ. Estimate

$$A > \int_{\mathbb{R}^n} |x|\mu_k = \int_{\mathbb{R}^n \setminus K_\epsilon} |x|\mu_k + \int_{K_\epsilon} |x|\mu_k \geq \int_{\mathbb{R}^n \setminus K_\epsilon} |x|\mu_k \geq A/\epsilon \int_{\mathbb{R}^n \setminus K_\epsilon} \mu_k \tag{5.60}$$

and thus

$$\int_{\mathbb{R}^n \setminus K_\epsilon} \mu_k < \epsilon. \tag{5.61}$$

Thus the sequence is tight and the convergent subsequence is guaranteed by Prokhorov's theorem. □

We will also need some continuity result on the free energy functional to ensure that a minimizing sequence can get close to the minimum. We begin with the entropy functional.

Proposition 22. *The entropy functional $D(\cdot)$ is lower semi-continuous with respect to weak convergence of measures on the space of probability measures with a bound on the average absolute deviation, i.e, fixing a constant A,*

$$\int |x|\mu < A \ \forall \mu. \tag{5.62}$$

We postpone the proof which will follow directly from that the Kullback-Leibler divergence is in fact the Legendre transform of another functional.

Lemma 3. *We have for probability measures μ and μ_0 where μ is absolutely continuous with respect to μ_0 that*

$$D(\mu||\mu_0) = \sup_{u \in \mathcal{C}_b(\mathbb{R}^n)} \int u\mu - \log \int e^u \mu_0. \tag{5.63}$$

Proof. By Jensen's inequality

$$\log \int e^u \mu_0 = \log \int e^u \frac{\mu_0}{\mu} \mu = \log \int e^{u + \log(\mu_0/\mu)} \mu \tag{5.64}$$

$$\leq \int u + \log(\mu_0/\mu)\mu = \int u\mu - \int \log(\mu/\mu_0)\mu \tag{5.65}$$

and thus after rearranging

$$D(\mu||\mu_0) \leq \int u\mu - \log \int e^u \mu_0. \tag{5.66}$$

To achieve equality we would like to pick $u = \log(\mu/\mu_0)$ in the supremum. But u such chosen might very well not be bounded and continuous. However using a standard approximation argument from measure theory, bounded continuous functions are dense in $L^1(\mathbb{R}^n)$ and thus we can take a sequence achieving equality in the supremum.

□

Proposition 22. Pick a probability measure μ_0 absolutely continuous with respect to Lebesgue with a non-zero density which is rapidly enough decaying such that $\int \log(\mu_0/\mathrm{d}x)\mu > -\infty$ for all μ satisfying the bound on the average absolute deviation. The measure with density proportional to $e^{-||x||}$ would do for example. Since μ_0 is non-zero μ is absolutely continuous with respect to μ_0 as well.

By the Lemma we have

$$D(\mu||\mu_0) = \sup_{u \in \mathcal{C}_b(\mathbb{R}^n)} \int u\mu - \log \int e^u \mu_0 \tag{5.67}$$

and since

$$D(\mu||\mu_0) = \int \log(\mu/\mu_0)\mu = \int \log(\mu/\mathrm{d}x)\mu - \int \log(\mu_0/\mathrm{d}x)\mu \tag{5.68}$$

we get

$$D(\mu) = \sup_{u \in \mathcal{C}_b(\mathbb{R}^n)} \int u\mu - \log \int e^u \mathrm{d}x + \int \log(\mu_0/\mathrm{d}x)\mu. \tag{5.69}$$

Clearly both $\int \log(\mu_0/\mathrm{d}x)\mu$ and $\int u\mu$ for $u \in \mathcal{C}_b(\mathbb{R}^\ltimes)$ are continuous with respect to the weak topology. It is a general fact that a supremum of continuous functionals is itself lower semi-continuous. Indeed, take a sequence $\{\mu_k\}_k$ of probability measures converging weakly to μ and estimate

$$\begin{aligned}
\liminf_k D(\mu_k) &= \sup_k \inf_{l>k} \sup_{u \in \mathcal{C}_b(\mathbb{R}^n)} \int u\mu_l - \log \int e^u \mathrm{d}x && (5.70)\\
&\geq \sup_k \sup_{u \in \mathcal{C}_b(\mathbb{R}^n)} \inf_{l>k} \int u\mu_l - \log \int e^u \mathrm{d}x && (5.71)\\
&= \sup_{u \in \mathcal{C}_b(\mathbb{R}^n)} \sup_k \inf_{l>k} \int u\mu_l - \log \int e^u \mathrm{d}x && (5.72)\\
&= \sup_{u \in \mathcal{C}_b(\mathbb{R}^n)} \int u\mu - \log \int e^u \mathrm{d}x && (5.73)\\
&= D(\mu). && (5.74)
\end{aligned}$$

Thus in total $D(\mu)$ is lower semi-continuous.

□

We also need a continuity result for the energy functional. This is more difficult and we only present a partial result in this direction. In particular given that we can restrict ourselves to measures with finite $1+\sigma$-absolute moment for σ positive we do have continuity.

Proposition 23. *Fix $\sigma > 0$. The energy functional $E(\cdot)$ is continuous with respect to the weak topology on the space of probability measures on $\mathbb{R}^n$ with a bound on the $1+\sigma$-absolute moment, i.e, all probability measures such that*

$$\int |x|^{1+\sigma}\mu < M \tag{5.75}$$

for some $M \in \mathbb{R}$.

Proof. To begin with, we can wonder what could be the problem. Recall the following form of energy functional

$$E(\mu) = \sup_T \int \langle x, T(x)\rangle \mu. \tag{5.76}$$

Where the sup is over transport maps T. As we have already seen, a supremum over continuous functionals is lower semi-continuous. Thus it suffices to prove that $E_T(\mu) = \int \langle x, T(x)\rangle$ is continuous. But $x \mapsto \langle x, T(x)\rangle$ need not be continuous nor bounded thus continuity of E_T is not direct. For the problem with boundedness we can use the bound on the moment and proceed similarly as to the proof of tightness. Denote by R a real number such that the polytope P is a subset of the open ball of radius R centered at the origin. Pick a continuous transport map T. Take $\epsilon > 0$. Pick a sequence $\{\mu_k\}_k$ of probability measures weakly converging to a probability measure μ such that

$$\int |x|^{1+\sigma}\mu_k < M \quad \forall k \tag{5.77}$$

for some $\sigma > 0$ and let $K_\epsilon = B(0, (4MR/\epsilon)^{1/\sigma})$ be the centered ball in $\mathbb{R}^n$ with radius $(4MR/\epsilon)^{1/\sigma}$. Estimate

$$|E(\mu_k) - E(\mu)| = |\sup_{T'} E_{T'}(\mu_k) - \sup_T E_T(\mu)| \tag{5.78}$$

$$\leq \sup_{T'} \left| \int_{\mathbb{R}^n \setminus K_\epsilon} \langle x, T'(x)\rangle \mu_k \right| + \sup_T \left| \int_{\mathbb{R}^n \setminus K_\epsilon} \langle x, T(x)\rangle \mu \right| + \tag{5.79}$$

$$\left| \sup_{T'} \int_{K_\epsilon} \langle x, T'(x)\rangle \mu_k + \sup_T \int_{K_\epsilon} \langle x, T(x)\rangle \mu_k \right|. \tag{5.80}$$

The last term being small is equivalent to proving continuity of the energy functional under the extra assumption that the measures all have compact support. This we postpone til the end of the proof. So for k large enough the last term is smaller than $\epsilon/2$. The first two terms we can treat in precisely the same way. Estimate

$$MR > R\int |x|^{1+\sigma}\mu > \int |x|^\sigma\langle x, T(x)\rangle\mu \tag{5.81}$$
$$= \int_{\mathbb{R}^n\setminus K_\epsilon} |x|^\sigma\langle x, T(x)\rangle\mu + \int_{\mathbb{R}^n} |x|^\sigma\langle x, T(x)\rangle\mu \tag{5.82}$$
$$\leq \int_{\mathbb{R}^n\setminus K_\epsilon} |x|^\sigma\langle x, T(x)\rangle\mu \leq 4MR/\epsilon \int_{\mathbb{R}^n\setminus K_\epsilon} \langle x, T(x)\rangle\mu. \tag{5.83}$$

And thus

$$\int_{\mathbb{R}^n\setminus K_\epsilon} \langle x, T(x)\rangle\mu \leq \epsilon/4. \tag{5.84}$$

Putting it all together we get

$$|E(\mu_k) - E(\mu)| < \epsilon \tag{5.85}$$

but ϵ was arbitrary so $E_T(\cdot)$ is continuous.

So why is $E(\mu)$ continuous for compactly supported measures? Here we will not be as detailed as it is rather involved. Assume μ also had compact support. Here it turns out to be better to work with the Kantorovich problem since then the problem is well-formulated for arbitrary measures. Write

$$E(\mu) = -\inf_\gamma \int -\langle x, y\rangle\gamma = -\inf_\gamma \int ||x-y||^2\gamma + 2\int ||x||^2\mu + 2\int ||y||^2\nu \tag{5.86}$$

as we have done before. The compact support ensures that the variances are finite. Also $\mu \mapsto \int ||x||^2\mu$ is manifestly continuous. The first functional is the cost of optimal transport between μ and ν with squared Euclidean cost. This turns out to be a distance on the space of probability measures on $\mathbb{R}^n$. More so, this distance metrizes the weak topology, a proof of which can be found in [12, Theorem 6.8]. It is a general fact which follow almost form definitions that any metric is doubly a continuous function with respect to the metric topology. □

Next we will prove a properness type-result for the energy functional in the one-dimensional case, a higher dimensional version is in preparation. There are some facts in one dimension which simplifies things. One of them is that the transport map is always monotone. This can be seen either directly from the minimization problem or from the fact that the transport map is realized as the gradient of a convex function. Another is that the absolute value function $|\cdot|$ is significantly simpler than in higher dimensions. Using these facts we prove:

Proposition 24. *Let μ be a probability measure on $\mathbb{R}$ absolutely continuous with respect to Lebesgue and with finite average absolute deviation. Let $P = [-2C, 2C]$ be an interval (polytope in one dimension) with $C > 0$, then after perhaps translating μ*

$$E(\mu) \geq \frac{C}{2} a \tag{5.87}$$

where a is the absolute average deviation of the translated measure.

Proof. We can work with the Monge-problem and assume the existence of a almost everywhere defined transport map T since μ is absolutely continuous and has finite average absolute deviation by Theorem 6. We translate μ until $T(x)x \geq 0$ where $T : \mathbb{R} \to \mathbb{R}$ is the transport map optimally transporting μ to $1_P \mathrm{d}p$. This works since 0 is an interior point of P, indeed, T changes sign somewhere and must do so in the interior of the convex hull of the support of μ, then just translate μ until the sign change of T occurs precisely at $x = 0$. Now estimate

$$2E(\mu) = 2\int T(x)x\mu = 2\int_{|T(x)|\geq C} T(x)x\mu + 2\int_{|T(x)|<C} T(x)x\mu \tag{5.88}$$

$$\geq 2\int_{|T(x)|\geq C} T(x)x\mu + 0 = 2\int_{|T(x)|\geq C} |T(x)||x|\mu \tag{5.89}$$

$$\geq 2C\int_{|T(x)|\geq C} |x|\mu \tag{5.90}$$

$$\geq C\int_{|T(x)|\geq C} |x|\mu + C\int_{|T(x)|<C} |x|\mu \tag{5.91}$$

$$= C\int |x|\mu = Ca \tag{5.92}$$

where in the third step we have used that $T(0) = 0$ and that T is monotone so that $T(x)x \geq 0$. In the fifth step we have used that the measure of $\{|T(x)| \geq C\}$ and

$\{|T(x)| < C\}$ is the same under μ, otherwise T would not transport μ to P, also that T is again monotone. By dividing by C we get the result.

□

5.1 A conditional proof of existence

What we so far have developed for the variational problem is unfortunately not enough for a complete existence proof of the variational problem or the weak toric Kähler–Einstein equation. What we will do nevertheless for explanatory purposes, is to sketch what an existence proof would look like using the theory developed for the free energy functional but also some extra conjectural facts. This will shed light on why the propositions stated and proven so far is of interest. The aim is to be as clear as possible as to when extra assumptions are used and what parts are precise.

In the following conjectural theorem that regards existence of a minimizer of the free energy, we will need a higher dimensional generalization of the properness-type-result in Proposition 24. Additionally, we need continuity of the energy functional. Remember, we only proved continuity of the energy functional in Proposition 23 in the class of probability measure with bounded $1+\epsilon$-moment for some positive ϵ. We will not be able to ensure such a bound on the minimizing sequence, only for $\epsilon = 0$ which is not enough.

Theorem (Conditional) 1. *Let P be a polytope in $\mathbb{R}^n$ with mass center at the origin. Then there is a probability measure with finite absolute average deviation minimizing the free energy functional.*

Proof. The only if is done by Proposition 18. For the if part, denote

$$\tilde{F} = \inf_{\mu} F(\mu) \tag{5.93}$$

where the infimum is over all probability measures on $\mathbb{R}$. By Proposition 14 for any measure with finite average absolute deviation the energy is finite and among those there are plenty of distributions with also finite entropy so $\tilde{F}$ is certainly not $+\infty$. Let $\{\mu_k\}_k$ be a sequence of probability measures such that $F(\mu_k) \to \tilde{F}$. We can assume that they are all absolutely continuous since otherwise we would have $\tilde{F} = +\infty$. Using Proposition 24 in one dimension, and a conjectural generalization in higher dimensions, and Proposition 20 we have the bound

$$F(\mu_k) \geq -\log(a_{\mu_k}) + \frac{C}{2}a_{\mu_k} + K_1 \tag{5.94}$$

where a_μ is the average absolute deviation of μ and K_1 is the constant independent of μ from Proposition 20 and $C \geq 0$ is the constant from Proposition 24 also independent of μ. From this bound we conclude that the infimum in (5.93) is finite since a linear function grows faster than the logarithm. Since the bound goes to infinity as a_μ goes to infinity we can assume without loss of generality that $\{\mu_k\}_k$ has a uniform bound on the average absolute deviation. By Proposition 21 we can find a subsequence $\{\mu_{k_l}\}_l$ which converges weakly to a probability measure μ with finite average absolute deviation.

At this point we would like to argue why we can further restrict the minimizing sequence to have a uniform bound also on the $1+\epsilon$-moment for some $\epsilon > 0$. Let that be conjectural. By Proposition 23 the energy functional is continuous on the minimizing sequence, and by Proposition 22 the entropy functional is lower semi-continuous. Thus the free energy functional is lower semi-continuous along the minimizing sequence and

$$F(\mu) \leq \liminf_k F(\mu_k) = \tilde{F} \tag{5.95}$$

so μ really is a minimizing probability measure of the free energy functional. □

Next one would like to say that μ has a density satisfying the toric Kähler–Einstein equation. Firstly, for it to satisfy the weak toric Kähler–Einstein equation, we simply need it to be an extremum of the free energy functional and satisfy $dF|_\mu = 0$. Then as we explained in the beginning of the section μ will satisfy the transport problem related to (4.51), which is by our definition a weak solution. But there is an important detail keeping us from concluding this right away. We would like to say that if $dF|_\mu \neq 0$ then we could just alter μ slightly and get a probability measure with lower free energy. But if the density of μ happens to be 0 somewhere then the perturbed measure need not be a positive measure any more. One possible solution might be to mirror the solution to the related problem in the variational problem varying ϕ studied in [6]. Anyhow this problem is a subject of future work.

Lastly one might want to investigate uniqueness. This part is largely unfinished but we present a possible future direction.

The problem is that one can quite easily see that along convex combinations, that is curves between two probability measures μ and ν given by $t \mapsto (1-t)\mu + t\nu$, the entropy is convex while the energy is concave so it is difficult to get a convexity result. However we can consider other classes of curves, namely *Wasserstein geodesics*. The *Wasserstein distance* $d(\mu, \nu)$ between two probability measures $\mu,\ \nu$ is simply the optimal cost of transporting one to the other but with cost function $c(x, y) = ||x-y||^2$. i.e,

$$d(\mu, \nu) = \inf_{\gamma \in \Pi(\mu,\nu)} \int ||x-y||^2 \gamma. \tag{5.96}$$

As we have mentioned before this is actually a metric on the space of probability measure metrizing the weak topology. Moreover this metrics space is geodesic, meaning that between any two probability measures μ, ν with finite second moments there is a curve ρ_t such that

$$d(\rho_t, \rho_s) = |t-s| d(\mu, \nu). \tag{5.97}$$

Thus if the energy functional was defined with the Euclidean cost then it would vary in this affine way along geodesics. By a result [13, Proposition 2.1] the entropy is actually convex along Wasserstein geodesics, and strictly convex if the two measures are not related by translation. Thus if our cost was Euclidean and we could assume finite second moments then the free energy would be convex along this class of curves, and after fixing the translational invariance strictly convex, sufficient for ensuring uniqueness of a minima. Additionally we would get uniqueness of the weak Kähler–Einstein equation since no other extrema could exist. However the cost $\langle x, y\rangle$, being intimately related to the Euclidean cost is not by any means equivalent to it and so this is only a possible direction.

5.2 The free energy as an invariant

In this section we leave the application of the variational problem to existence and instead study the free energy functional in its own regard. In defining the free energy functional we have only used biholomorphically invariant data. Additionally the Kähler–Einstein equation itself is actually biholomorphically invariant as can be seen from a simple calculation. We have not talked much about uniqueness of Kähler–Einstein metrics but they are indeed unique when they exist up to a scalar multiple.

This scalar disappears when we get to the Kähler–Einstein measure μ which we demand to be a probability measure. Thus the free energy functional evaluated at the Kähler–Einstein measure is an invariant of the toric complex manifold. For reasons not dwelled on here, we will involve the volume of the polytope as well and consider for a Fano compact complex toric manifold X with fundamental polytope P and Kähler–Einstein measure μ_X the invariant $\mathcal{F}_X$ when μ_X exists is defined to be

$$\mathcal{F}_X = \mathrm{Vol}(P)F(\mu_X). \tag{5.98}$$

To calculate this invariant for a given X, one needs to calculate the volume, the energy and entropy of the Kähler–Einstein measure. With the following theorem, computing the energy is actually superfluous. The proof will depend heavily on the not yet proven fact that μ_X really is a unique minimizer of the free energy functional for $\beta = 1$.

Theorem (Conditional) 2. *The energy functional evaluated at the Kähler–Einstein measure of a compact toric complex manifold X is always given by the dimension. i.e,*

$$E(\mu_X) = \dim(X). \tag{5.99}$$

Proof. Recall the free energy functional (5.3) with the inverse temperature β included. Define

$$G_\beta(\mu) := \beta F(\mu) = \beta E(\mu) + D(\mu). \tag{5.100}$$

Define further

$$g(\beta) = \inf_{\mu \in \mathcal{P}} G_\beta(\mu). \tag{5.101}$$

Claim that $\frac{\mathrm{d}}{\mathrm{d}\beta} g(\beta)|_{\beta=1} = E(\mu_X)$ where μ_X is the Kähler–Einstein metric on X.

Introducing $t = \beta - 1$ note that

$$g(t) = \inf_{\mu \in \mathcal{P}} E(\mu) + D(\mu) + tE(\mu). \tag{5.102}$$

To proceed we must use that the infimum is attained by a unique measure μ_β for every β in a neighborhood of 1, something which would follow from a existence theory like the one partly developed earlier in the section.

After that, one can use essentially the same proof as in the proof of Proposition 16 and conclude the claim. If we can compute $\frac{\mathrm{d}}{\mathrm{d}}g(\beta)|_\beta = 1$ we are done. Note that if ϕ is a solution to the critical point equation (5.27) with $\beta = 1$ then $\phi_\beta = \frac{1}{\beta}(\phi(\beta z)+C)$, for a correctly chosen constant C, is a solution to the Kähler–Einstein equation with β arbitrary positive real but with the same boundary condition. Indeed

$$\frac{1}{\mathrm{Vol}(P)} \det[\frac{\partial^2}{\partial x_i \partial x_j}\phi_\beta] = \frac{1}{\mathrm{Vol}(P)} \det\left[\frac{\partial^2}{\partial x_i \partial x_j}\left(\frac{1}{\beta}\phi(\beta z) + C\right)\right] = \frac{\beta^n}{\mathrm{Vol}(P)} \det[\frac{\partial^2}{\partial x_i \partial x_j}](\beta z) \tag{5.103}$$

$$= \beta^n e^{-\phi(\beta z)} = \beta^n e^C e^{-\beta\phi_\beta} = e^{-\beta\phi_\beta} \tag{5.104}$$

where we have chosen $C = -n\log(\beta)$. The associated transport map will be $T_\beta = \nabla\phi_\beta = (\nabla\phi)(\beta z)$ and thus ϕ_β also satisfies the boundary condition $\overline{\nabla\phi_\beta(\mathbb{R}^n)} = P$. Additionally $e^{-\beta\phi_\beta}$ has the correct normalization since

$$\int e^{-\beta\phi_\beta}\mathrm{d}x = \int e^{-\phi(\beta x)}\beta^n \mathrm{d}x = \int e^{-\phi(x')}\mathrm{d}x' = 1. \tag{5.105}$$

where $x' = \beta x$. We proceed by computing the energy of $\mu_\beta = e^{-\phi_\beta}$.

$$E(\mu_\beta) = -\int \langle x, T_\beta(x)\rangle \mu_\beta = -\beta^n \int \langle x, \nabla\phi_\beta(x)\rangle e^{-\phi(\beta x)}\mathrm{d}x \tag{5.106}$$

$$= -\beta^{(n-1)} \int \langle \beta x, (\nabla\phi)(\beta x)\rangle e^{-\phi(\beta x)}\mathrm{d}x \tag{5.107}$$

$$= -\frac{1}{\beta}\int \langle x', (\nabla\phi)(x')\rangle e^{-\phi(x')}\mathrm{d}x' \tag{5.108}$$

$$= \frac{1}{\beta}E(\mu_1) \tag{5.109}$$

and the entropy

$$D(\mu_\beta) = \int \log(e^{-\beta\phi_\beta})e^{-\beta\phi_\beta}\mathrm{d}x \tag{5.110}$$
$$= \log(\beta^n)\int \mu_\beta + \beta^n \int -\phi(\beta x)e^{-\phi(\beta x)}\mathrm{d}x \tag{5.111}$$
$$= n\log(\beta) + \int -\phi(x')e^{-\phi(x')}\mathrm{d}x' \tag{5.112}$$
$$= \frac{n}{2}\log(\beta) + D(\mu_1). \tag{5.113}$$

Thus

$$\frac{\mathrm{d}}{\mathrm{d}\beta}g(\beta)|_{\beta=1} = \frac{\mathrm{d}}{\mathrm{d}\beta}|_{\beta=1}E(\mu_1) + D(\mu_1) + n\log(\beta) = n \tag{5.114}$$

which completes the proof. □

So one needs only to compute the volume and the entropy to calculate $\mathcal{F}$.

6

Complex projective space

It is always useful to have an example of a mathematical object exemplifying a lot of the theory without being trivial. In our case the complex projective space plays precisely that role. It is a compact toric complex Fano variety with an explicit Kähler–Einstein metric.

Definition 16. *The Complex Projective space* $\mathbb{C}\mathrm{P}^n$ *is a complex manifold with underlying set* $\mathbb{C}^{n+1} \setminus \{0\}/\sim$ *where* $Z \sim W$ *if* $Z = \lambda W$ *for* $\lambda \in \mathbb{C}$*. We denote a point in* $\mathbb{C}\mathrm{P}^n$*, i.e, a equivalence class by* $[Z_0 : ... : Z_n]$*. The manifold structure is given by the following charts for* $i = 0, 1, ..., n$

$$\psi_i : \mathbb{C}\mathrm{P}^n \cap \{Z_i \neq 0\} \to \mathbb{C}^n \tag{6.1}$$

$$[Z_0 : Z_1 : ... : Z_n] \mapsto (Z_0/Z_i, ..., Z_{i-1}/Z_i, Z_{i+1}/Z_i, ..., Z_n/Z_i) \tag{6.2}$$

It is simple to verify that the transitions maps are biholomorphic.

$\mathbb{C}\mathrm{P}^n$ is an example of a toric complex manifold as can be seen by introducing the toric action given by multiplication in the n last components as

$$\lambda \cdot [Z_0 : .... : Z_1] = [Z_0 : \lambda Z_1 : ... : \lambda Z_n] \tag{6.3}$$

for $\lambda \in \mathbb{C}^{*n}$ and $[Z_0 : ... : Z_1]$.

The structure of complex projective space as a quotient space equips it with a natural line bundle.

Definition 17. *Define the tautological line bundle* $\mathcal{O}(-1)$ *by*

$$\pi : \mathrm{Bl}_0\mathbb{C}^{n+1} \to \mathbb{C}\mathrm{P}^n \tag{6.4}$$
$$((Z'_0, ..., Z'_n), [Z_0 :, ..., : Z_n]) \mapsto [Z_0 :, ..., : Z_n]. \tag{6.5}$$

where $\mathrm{Bl}_0\mathbb{C}^{n+1}$ *is the blow-up of* $\mathbb{C}^{n+1}$ *is the blowup of* $\mathbb{C}^{n+1}$ *at the origin. The blow-up is constructed as a certain subset of* $\mathbb{C}^{n+1} \times \mathbb{C}\mathrm{P}^n$ *which is the "diagonal" everywhere except at the origin. At the origin there is a whole copy of* $\mathbb{C}\mathrm{P}^n$*. The construction can be shown to be a complex manifold. Intuitively the tautological line bundle is constructed by adding the complex line represented by each point of* $\mathbb{C}\mathrm{P}^n$*, to* $\mathbb{C}\mathrm{P}^n$*.*

We can also construct a second line bundle by simply taking the dual $\mathcal{O}(-1)^*$ and we call this the *hyperplane bundle* and denote it by $\mathcal{O}(1)$. There is a nice characterization of the space of holomorphic sections $H^0(\mathbb{C}\mathrm{P}^n, \mathcal{O}(1))$ of $\mathcal{O}(1)$. Let $s \in H^0(\mathbb{C}\mathrm{P}^n, \mathcal{O}(1)$. To evaluate s we need both a point $[Z_0 : ... : Z_n]$ in $\mathbb{C}\mathrm{P}^n$ and a point in the line above $[Z_0 : ... : Z_n]$. Thus we can choose to view s as a function on the total space of $\mathcal{O}(-1)$. But s should still be linear on the fibers and thus one-homogeneous on $\mathbb{C}^{n+1} \setminus \{0\}$. s must also have a holomorphic dependence on the point on $\mathbb{C}\mathrm{P}^n$ and thus s has to be a one-homogeneous holomorphic function on $\mathbb{C}^{n+1} \setminus \{0\}$. But by essentially the same arguments any one-homogeneous holomorphic function on $\mathbb{C}^{n+1} \setminus \{0\}$ can be seen as a section on $H^0(\mathbb{C}\mathrm{P}^n, \mathcal{O}(1))$. Thus we have with this correspondence in mind

$$H^0(\mathbb{C}\mathrm{P}^n, \mathcal{O}(1)) = \mathrm{span}\{Z_i | i = 0, ..., n\}, \tag{6.6}$$

i.e, homogeneous polynomials in $n + 1$ variables of order 1.

A natural question to ask is if the natural bundles $\mathcal{O}(-1)$ or $\mathcal{O}(1)$ on $\mathbb{C}\mathrm{P}^n$ originating from the quotient structure is related to the canonical bundle $K_{\mathbb{C}\mathrm{P}^n}$. To understand how they are connected, we need to introduce tensor powers and duals of lines bundles. Essentially, any vector space operation makes sense on a vector bundle by applying the operation fiberwise. The trivializations are easily lifted to the tensor powered or dualised bundles. One can see that the corresponding transition maps, as defined in section chapter 2, of a tensor product of line bundles corresponds to a product of the transition maps. Likewise the dual of a line bundle will have transition maps that are the reciprocal of the transition maps to the original line bundle. We will adapt the notation $\mathcal{O}(n)$ for the n'th tensor power of $\mathcal{O}(1)$ with itself. We quickly note that the sections of tensor powers $\mathcal{O}(k)$ corresponds to the k-homogeneous functions on $\mathbb{C}^{n+1} \setminus \{0\}$ for the same reason that sections of $\mathcal{O}(1)$ corresponded to the one-homogeneous functions on $\mathbb{C}^{n+1} \setminus \{0\}$. With this correspondence in mind

$$H^0(\mathbb{C}P^n, \mathcal{O}(k))) = \{\text{homogeneous polynomials of degree } k \text{ in the variables } Z_0, ..., Z_n\}. \tag{6.7}$$

Going back to our question we formulate and prove the following proposition.

Proposition 25. *There an isomorphism of complex line bundles*

$$K^*_{\mathbb{C}\mathrm{P}^n} \cong \mathcal{O}(n+1). \tag{6.8}$$

Proof. We have a cover of $\mathbb{C}\mathrm{P}^n$ of charts of the form $Z_i \neq 0$ for $i = 0, ..., n$. In these charts we can trivialize $\mathcal{O}(1)$ with the corresponding homogeneous coordinates Z_i $\forall i$. Fix i and j and lets compute the transition map between the trivializations Z_i and Z_j in the natural coordinates z_k on $Z_j \neq 0$. Recall that we have $(z_1, ..., z_n) \mapsto [z_1 : ...z_j : 1 : z_{j+1} : ... : z_n]$.

In these coordinates we have

$$\Psi^{\mathcal{O}(1)}_{ji} = Z_i/Z_j = \begin{cases} z_{i+1}, & i < j \\ z_i, & i > j. \end{cases} \tag{6.9}$$

Note here that the coordinates z_k used also depend themselves on j; they are always the natural coordinates on $\{Z_j \neq 0\}$ but we have suppressed this dependence for brevity. The corresponding transition functions for $\mathcal{O}(n+1)$ are

$$\Psi^{\mathcal{O}(n+1)}_{ji} = \begin{cases} z^{n+1}_{i+1}, & i < j \\ z^{n+1}_i, & i > j \end{cases} \tag{6.10}$$

Next we observe that we can also trivialize $K_{\mathbb{C}\mathrm{P}^n}$ over the same open cover given by

$\{Z_i \neq 0\}_i$. We use the trivializing section

$$dw_1 \wedge \dots \wedge dw_n. \tag{6.11}$$

over $\{Z_i \neq 0\}$ and the trivializing section

$$dz_1 \wedge \dots \wedge dz_n \tag{6.12}$$

over $\{Z_j \neq 0\}$. To express the transition function in the z_k-coordinates we have to change coordinates on the differential form $dw_1 \wedge \dots \wedge dw_n$. For $j < i$ the coordinate change is given by

$$w_k = \begin{cases} \frac{Z_{k-1}}{Z_i} & k \leq i \\ \frac{Z_k}{Z_i} & k > i \end{cases} = \begin{cases} \frac{Z_{k-1}}{Z_j}\frac{Z_j}{Z_i} & k \leq i \\ \frac{Z_k}{Z_j}\frac{Z_j}{Z_i} & k > i \end{cases} = \begin{cases} \frac{z_k}{z_j} & k \leq j \\ \frac{1}{z_i} & k = j+1 \\ \frac{z_{k-1}}{z_i} & j+1 < k \leq i \\ \frac{z_k}{z_i} & i < k. \end{cases} \tag{6.13}$$

The Jacobian of this transformation is mostly diagonal and the $(j{+}1)$'th row contains only one non-zero element. Expanding over this row

$$\det J[w](z) = \frac{(-1)^{i+j}}{z_i^{n+1}}. \tag{6.14}$$

and thus

$$dw_1 \wedge \dots \wedge dw_n = -\frac{(-1)^{i+j}}{z_i^{n+1}} dz_1 \wedge \dots \wedge dz_n. \tag{6.15}$$

Thus we have computed the transition function

$$\Psi_{ji}^{K_{\mathbb{C}P^n}} = \frac{(-1)^{i+j}}{z_i^{n+1}} \tag{6.16}$$

whenever $j < i$. For $j > i$ the coordinate change is instead

$$w_k = \begin{cases} \frac{Z_{k-1}}{Z_i} & k \leq i \\ \frac{Z_k}{Z_i} & k > i \end{cases} = \begin{cases} \frac{Z_{k-1}}{Z_j}\frac{Z_j}{Z_i} & k \leq i \\ \frac{Z_k}{Z_j}\frac{Z_j}{Z_i} & k > i \end{cases} = \begin{cases} \frac{z_k}{z_{i+1}} & k \leq j \\ \frac{z_{k+1}}{z_{i+1}} & k = j+1 \\ \frac{1}{z_{i+1}} & k = j \\ \frac{z_k}{z_{i+1}} & k > j. \end{cases} . \tag{6.17}$$

The Jacobian is again mostly diagonal and expanding over the j:th row one has

$$\det J[w](z) = \frac{(-1)^{i+j}}{z_{i+1}^{n+1}} \tag{6.18}$$

and thus

$$\mathrm{d}w_1 \wedge ... \wedge \mathrm{d}w_n = \frac{(-1)^{i+j}}{z_i^{n+1}} \mathrm{d}z_1 \wedge ... \wedge \mathrm{d}z_n. \tag{6.19}$$

In total we have computed all relevant transition functions for $K_{\mathbb{CP}^n}$ and they are

$$\tilde{\Psi}_{ji} = \begin{cases} \frac{(-1)^{i+j}}{z_{i+1}^{n+1}} & i < j \\ \frac{(-1)^{i+j}}{z_i^{n+1}} & i > j \end{cases} . \tag{6.20}$$

The dual of the canonical bundle $K^*_{\mathbb{CP}^n}$ will then have transition functions

$$\Psi'_{ji} = \begin{cases} (-1)^{i+j} z_{i+1}^{n+1} & i < j \\ (-1)^{i+j} z_i^{n+1} & i > j \end{cases} . \tag{6.21}$$

But comparing with the transition functions for $\mathcal{O}(n+1)$ we observe that they are equivalent. Indeed, take a holomorphic re-trivialization given by $f_k(z) = (-1)^k$ subject to the cover. Then

$$\frac{f_i}{f_j}\Psi_{ji} = \tilde{\Psi}_{ji} \tag{6.22}$$

and we conclude the isomorphism.

□

We continue by investigating the toric structure of $\mathbb{CP}^n$. Consider the torus action

$$\lambda \cdot [Z_0 : ... : Z_n] = [Z_0 : \lambda_1 Z_1 : \lambda_2 Z_2 : ... : \lambda_n Z_n] \tag{6.23}$$

for $[Z_0 : ... : Z_n] \in \mathbb{CP}^n$ and $\lambda = (\lambda_1, ..., \lambda_n) \in \mathbb{C}^{*n}$. It is evident that the set $\{Z_i \neq 0 \; \forall Z_i\} \in \mathbb{CP}^n$ is a dense, open, free orbit and thus $\mathbb{CP}^n$ is a toric compact complex manifold. There are plenty of toric line bundles on $\mathbb{CP}^n$ with respect to this action. All $\mathcal{O}(k)$ $\forall k$ for example have a lifted action which acts precisely as (6.23) extended symbolically to the sections of $\mathcal{O}(k)$, recalling that the sections are just order k homogeneous polynomials. But there are more actions on $\mathcal{O}(k)$ that when projected down to the base space is the same as the one we defined. We can twist the action also multiplying all sections with a *twisting function*, holomorphic in λ. Clearly the projected action on $\mathbb{CP}^n$ stays the same. If we consider $\mathcal{O}(n+1)$ with twisting function (with respect to the natural action) given by $f(\lambda) = \lambda_1^{-1}\lambda_2^{-1}...\lambda_n^{-1}$ then we have that under the isomorphism in Proposition 25 the action on $\mathcal{O}(n+1)$ is precisely the same as the natural action on K_X^*. Indeed we know from the proof of the proposition that the section Z_0^{n+1} of $\mathcal{O}(n+1)$ corresponds, perhaps up to a sign, to the section $\frac{\partial}{\partial z_1} \wedge ... \wedge \frac{\partial}{\partial z_n}$ of K_X in the natural coordinate system $\{z_i\}_i$ on $Z_0 \neq 0$. With the chosen twisting function, the action on $\mathcal{O}(n+1)$, Z_0^{n+1} transforms under the action as $\lambda \cdot Z_0^{n+1} = \lambda_1^{-1}\lambda_2^{-1}...\lambda_n^{-1} Z_0^{n+1}$ since Z_0 is invariant under the original action on $O(1)$. This is precisely how $\frac{\partial}{\partial z_1} \wedge ... \wedge \frac{\partial}{\partial z_n}$ transforms under the action on $K_{\mathbb{CP}^n}^*$.

With this, it is an easy task to enumerate the eigenstates of $H^0(\mathbb{CP}^n, K_{\mathbb{CP}^n})$ under the torus action. The eigenstates are simply the monomials in $H^0(\mathbb{CP}^n, \mathcal{O}(k+1)$, i.e, the degree $n+1$ monomials in the variables $Z_0, ..., Z_n$. The character of such a monomial $Z_0^{p_0} Z_1^{p_1}$ under the twisted action is given by

$$\lambda_1^{p_1-1}\lambda_2^{p_2-1}...\lambda_n^{p_n-1}. \tag{6.24}$$

The integral vector $p = (p_1, ..., p_n)$ is constrained by having positive entries, and a sum less or equal to $n+1$ and thus lying in the polytope given by the positive quadrant in $\mathbb{R}^n$ intersected by the half plane $\{x \in \mathbb{R}^n | (1, 1, ..., 1) \cdot x \leq n+1\}$. The required twisting of the action translates this polytope by the vector $(-1,-1,...,-1)$ and the resulting polytope is the fundamental polytope of $(\mathbb{CP}^n, K^*_{\mathbb{CP}^n})$. The polytope can be defined by the inequalities

$$(-1,-1,...,-1)\cdot x \geq -1 \tag{6.25}$$
$$(1,0,0,...,0)\cdot x \geq -1 \tag{6.26}$$
$$(1,1,0,...,0)\cdot x \geq -1 \tag{6.27}$$
$$\vdots \tag{6.28}$$
$$(0,0,...,1)\cdot x \geq -1. \tag{6.29}$$

thus we have proven that $\mathbb{CP}^n$ is Fano if we use the converse to Theorem 2 which is sketched under the proof. But there are definitely easier ways of seeing this, for example by considering the Kähler–Einstein metric on it which we will construct. One can also perform the slightly bothersome computation of computing the barycenter of this polytope and find that it is indeed the origin. This would be in line with what is described in the section on the variational problem where it was argued, but also cited, that the condition that the mass center is the origin is a necessary and sufficient condition for the existence of a Kähler–Einstein metric.

We continue by constructing the Kähler–Einstein metric on $\mathbb{CP}^n$. We will find it in a rather ad hoc way by first constructing a natural metric on $\mathcal{O}(-1)$ and then consider the repeated tensor product of the dual metric on $\mathcal{O}(n+1) = K^*_X$ and show that it is Kähler–Einstein.

Consider the Euclidean metric in homogeneous coordinates,

$$||Z|| = \sqrt{|Z_0|^2 + |Z_1|^2 + ... + |Z_n|^2}. \tag{6.30}$$

It is a positive, positively one-homogeneous function on the total space of $\mathcal{O}(-1)$. Thus it can be seen as a metric on $\mathcal{O}(-1)$, it varies smoothly and acts just as a one-dimensional complex vector space metric on every fiber (complex line). We denote this metric by $||\cdot||_{\mathcal{O}(-1)}$. Consider the dual metric $||\cdot||_{\mathcal{O}(1)}$ on $\mathcal{O}(1)$ and take the trivializing section Z_0 of $\mathcal{O}(1)$ over $\{Z_0 \neq 0\} \to \mathbb{CP}^n$. We would like to compute the length $||Z_0||_{\mathcal{O}(1)}$ in order to compute the weight function. By definition of the dual metric we have

$$||Z_0||_{\mathcal{O}(1)} = ||Z_0^*||_{\mathcal{O}(-1)} \tag{6.31}$$

where Z_0^* is the unique section satisfying

$$\langle Z_0, Z_0^* \rangle = ||Z_0||_{\mathcal{O}(1)}. \tag{6.32}$$

Thus we must have

$$Z_0^* : \mathbb{C}\mathrm{P}^n \to \mathcal{O}(-1) \tag{6.33}$$

$$[Z_0 : Z_1 : ... : Z_n] \mapsto ||Z_0||^2_{\mathcal{O}(1)}(1, Z_1/Z_0, ..., Z_n/Z_0). \tag{6.34}$$

Using this we get

$$||Z_0||_{\mathcal{O}(1)} = ||Z_0^*||_{\mathcal{O}(-1)} = ||Z_0||^2_{\mathcal{O}(1)} \sqrt{1 + \left|\frac{Z_1}{Z_0}\right|^2 + ... + \left|\frac{Z_n}{Z_0}\right|^2} \tag{6.35}$$

and consequently

$$||Z_0||^2_{\mathcal{O}(1)} = \left(1 + \left|\frac{Z_1}{Z_0}\right|^2 + ... + \left|\frac{Z_n}{Z_0}\right|^2\right)^{-1/2} \tag{6.36}$$

thus the weight of $|| \cdot ||_{\mathcal{O}(-1)}$ in this trivialization is

$$\phi = -\log ||Z_0||^2_{\mathcal{O}(1)} = \log(1 + \left|\frac{Z_1}{Z_0}\right|^2 + ... + \left|\frac{Z_n}{Z_0}\right|^2) = \log(1 + \sum_i |z_i|^2) \tag{6.37}$$

where $z_i,\ i = 1, .., n$ is the natural coordinate system on $\{Z_0 \neq 0\} \subset \mathbb{C}\mathrm{P}^n$. Taking the $n+1$:th tensor power of this metric with itself we find a metric on $O(n+1)$

whose weight function over $\{Z_0 \neq 0\}$ we will abusively denote with ϕ as well and it is

$$\phi = (n+1)\log(1 + \sum_i |z_i|^2). \tag{6.38}$$

We would like to compute the complex Monge-Ampere operator on ϕ appearing in the Kähler–Einstein equation. The direct computation is difficult due to the determinant appearing and thus we will present a more indirect approach. We will use the connection with $\mathbb{C}\mathrm{P}^n$ to consider $\mathbb{C}^{n+1}$ with variables $Z = (Z_0, ..., Z_n)$. Define

$$\Phi(Z) = \log(|Z|^2). \tag{6.39}$$

Clearly

$$(\partial\bar{\partial}e^{\Phi}(Z))^{n+1} = \mathrm{d}Z\mathrm{d}\bar{Z} \tag{6.40}$$

where $\mathrm{d}Z = \mathrm{d}Z_0 \wedge ... \wedge \mathrm{d}Z_n$. By the product rule

$$\partial\bar{\partial}e^{\Phi(Z)} = \partial(\bar{\partial}\Phi(Z)e^{\Phi(Z)}) = (\partial\bar{\partial}\Phi(Z) + \partial\Phi(Z) \wedge \bar{\partial}\Phi(Z))e^{\Phi(Z)} \tag{6.41}$$

and thus

$$(\partial\bar{\partial}e^{\Phi(Z)})^{n+1} = \Big((\partial\bar{\partial}\Phi(Z))^{n+1} + (n+1)(\partial\bar{\partial}\Phi(Z))^n \wedge \partial\Phi(Z) \wedge \bar{\partial}\Phi(Z)\Big)\, e^{(n+1)\Phi(Z)} \tag{6.42}$$

since all other terms contain wedges of more than one copy of $\partial\Phi(Z)$ and thus vanishes. Now note that $\Phi(Z)$ on $\mathbb{C}^{n+1} \setminus \{0\}$ enjoys the scaling property

$$\Phi(\eta Z) = \Phi(Z) + \log(|\eta|^2). \tag{6.43}$$

for $\eta \in \mathbb{C}$. Thus the matrix defined by the coefficients of the two-form $\partial\bar{\partial}\Phi(Z)$ has a 0 eigenvalue and the determinant appearing in the single coefficient in $(\partial\bar{\partial}\Phi(Z))^{n+1}$

vanishes. So far we have shown that

$$(n+1)(\partial\bar{\partial}\Phi(Z))^n \wedge \partial\Phi(Z) \wedge \bar{\partial}\Phi(Z) = e^{-(n+1)\Phi(Z)} \mathrm{d}Z \wedge \mathrm{d}\bar{Z}. \tag{6.44}$$

At this point we would like to bring in the coordinates on $\mathbb{C}^n$ so assume $Z_0 \neq 0$ and set $z_i = Z_i/Z_0$ for $i = 1, ..., n$ and $w = 2\log(Z_0)$ for some branch of log. We are doing this computation locally so any branch with branch cut away from Z_0 will do. Write $Z = e^{w/2}(1, z_1, z_2, ..., z_n)$. In these variables

$$\Phi(Z) = \log(\sum_i |Z_i|) = \log(|e^w|(1 + |z_1|^2 + |z_2|^2 + ... + |z_n|^2)) = \Re w + \frac{1}{n+1}\phi(z). \tag{6.45}$$

where $\Re w$ is the real part of w. Now note that

$$\partial\bar{\partial}\Re w = \partial\bar{\partial}\frac{1}{2}(w + \bar{w}) = \frac{1}{2}(\partial(\bar{\partial}w) + \partial\overline{\partial w} = 0 \tag{6.46}$$

since $\partial\bar{\partial}$ is invariant as an operator under change of variables and w is holomorphic. So $\partial\bar{\partial}\Phi(Z) = \partial\bar{\partial}\phi(z)$. The left hand side of (6.44) becomes

$$(\partial\bar{\partial}\phi(z)) \wedge \partial\left(\Re w + \frac{1}{n+1}\phi(z)\right) \wedge \bar{\partial}\left(\Re w + \frac{1}{n+1}\phi(z)\right) = \tag{6.47}$$

$$\frac{1}{4}(\partial\bar{\partial}\phi(z)) \wedge \mathrm{d}w \wedge \mathrm{d}\bar{w}. \tag{6.48}$$

For the right hand side we observe that the inverse change of coordinates is $Z_i = e^{w/2}Z_0$ and $Z_0 = e^{w/2}$. The Jacobian is mostly diagonal and after developing along the first row one finds

$$\det[Z](w, z) = \frac{1}{2}e^{(n+1)w/2}, \tag{6.49}$$

consequently

$$\mathrm{d}Z \wedge \mathrm{d}\bar{Z} = \frac{1}{4}e^{(n+1)\Re w}\mathrm{d}w \wedge \mathrm{d}\bar{w} \wedge \mathrm{d}z \wedge \mathrm{d}\bar{z}. \tag{6.50}$$

Using this the right hand side becomes

$$e^{-(n+1)\Phi(Z)}\mathrm{d}Z \wedge \mathrm{d}\bar{Z} = \frac{1}{4}e^{-(n+1)\Phi(Z)}e^{(n+1)\Re w}\mathrm{d}w \wedge \mathrm{d}w \wedge \mathrm{d}z \wedge \mathrm{d}\bar{z} \tag{6.51}$$

$$= \frac{1}{4}e^{-\phi(z)}\mathrm{d}w \wedge \mathrm{d}w \wedge \mathrm{d}z \wedge \mathrm{d}\bar{z}. \tag{6.52}$$

In total (6.44) becomes, after "cancelling" $\mathrm{d}w \wedge \mathrm{d}\bar{w}$

$$(\partial\bar{\partial}\phi(z)) = e^{-\phi(z)}\mathrm{d}z \wedge \mathrm{d}\bar{z} \tag{6.53}$$

which the Kähler–Einstein equation.

On the title page is a visualization of the Kähler–Einstein measure μ on $\mathbb{C}\mathrm{P}^2$. The egg-shaped structure is a contour plot of the Kähler–Einstein measure. The square grid is mapped to the skewed grid inside the polytope via the transport map $\nabla\phi$ for $\phi = -\log(\mu/\mathrm{d}x)$. Recall also how the measure μ is transported to the uniform measure on the polytope.

6.1 The Free energy of complex projective space

Complex projective space is the main example of compact complex manifolds with explicit Kähler–Einstein metrics. Other examples are products of complex projective spaces. Going back to the free energy functional, the strong connection between it and Kähler–Einstein metrics suggests that evaluating the free energy functional on a measure related to a Kähler–Einstein metric could be an interesting numerical invariant. With this in mind we calculate the entropy of the measure associated with the Fubini–Study metric on $\mathbb{C}P^n$. As we have seen the Fubini–Study metric can be represented by the Kähler potential

$$\phi = (n+1)\log(1 + \sum_i |z_i|^2) \tag{6.54}$$

where $z_i = Z_i/Z_0$ is the standard affine chart on $\mathbb{C}^n \hookrightarrow \mathbb{C}\mathrm{P}^n$. This is the standard way of representing the Fubini–Study metric, where the section $\mathrm{d}z_1 \wedge ... \wedge \mathrm{d}z_n$ is used as trivializing section. But as we have seen this section is not $T^n_{\mathbb{C}}$-invariant. Over $\mathbb{C}^{n*} \hookrightarrow \mathbb{C}^n$ we can use the same coordinates but we will use the $T^n_{\mathbb{C}}$-invariant section

$s_0 = \frac{1}{\prod_{i=1}^n z_i} \mathrm{d}z_1 \wedge ... \wedge \mathrm{d}z_n$. The weight representing Fubini–Study with respect to this trivializing section is

$$\phi = (n+1)\log(1 + \sum_i |z_i|^2) - \log(\prod_{i=1}^n |z_i|^2). \tag{6.55}$$

Introducing the Log-coordinates $x_i = \log|z_i|^2$ the associated measure is

$$\mu = \frac{1}{Z} \frac{\prod_{i=1}^n e^{x_i}}{(1 + \sum_{i=1}^n e^{x_i})^{n+1}} \mathrm{d}x \tag{6.56}$$

where Z is a normalizing factor. We perform the variable substitution

$$e^{x_i} = t_i \tag{6.57}$$

$$\prod_{i=1}^n e^{x_i} \mathrm{d}x = \mathrm{d}t \tag{6.58}$$

and get

$$Z = \int_{\mathbb{R}_+^n} \frac{1}{(1 + \sum_{i=1}^n t_i)^{n+1}} \mathrm{d}t. \tag{6.59}$$

Perform the variable substitution

$$\sum_{i=1}^n t_i = u_1 \tag{6.60}$$

$$t_k = u_k \text{ for } k > 1 \tag{6.61}$$

$$\mathrm{d}t = \mathrm{d}u \tag{6.62}$$

to get

$$Z = \int_{\mathbb{R}_+^{n-1}} \int_{\sum_{i=2}^n u}^{\infty} \frac{1}{(1+u_1)^{n+1}} \mathrm{d}u = \left[-\frac{1}{n}(1+u_1)^{-n} \right]_{u = \sum_{i=2}^n u}^{\infty} \tag{6.63}$$

$$= \frac{1}{n} \int_{\mathbb{R}_+^{n-1}} \frac{1}{(1 + \sum_{i=2}^n t_i)^n} \tag{6.64}$$

which we can identify with the expression (6.59) albeit with a different constant in front and with $n+1$ exchange to n. Following the procedure a total of $n-1$ times gives

$$\frac{1}{n!}\int_0^\infty \frac{1}{(1+t)^2}\mathrm{d}t = \frac{1}{n!} \tag{6.65}$$

so that $Z = 1/n!$. We continue with the entropy calculation.

$$D(\mu) = \int_{\mathbb{R}^n} \log(n!\frac{\prod_{i=1}^n e^{x_i}}{(1+\sum_{i=1}^n e^{x_i})^{n+1}})n!\frac{\prod_{i=1}^n e^{x_i}}{(1+\sum_{i=1}^n e^{x_i})^{n+1}}\mathrm{d}x. \tag{6.66}$$

Again it is useful to perform the substitution

$$e^{x_i} = t_i \tag{6.67}$$

$$\prod_{i=1}^n e^{x_i}\mathrm{d}x = \mathrm{d}t \tag{6.68}$$

to get

$$D(\mu) = \int_{\mathbb{R}_+^n} \log(N!\frac{\prod_{i=1}^n t_i}{(1+\sum_{i=1}^n t_i)^{n+1}})N!\frac{1}{(1+\sum_{i=1}^n t_i)^{n+1}}\mathrm{d}t \tag{6.69}$$

$$= \log(n!)n!\int_{\mathbb{R}_+^n} \frac{1}{(1+\sum_{i=1}^n t_i)^{n+1}}\mathrm{d}t + n!\int_{\mathbb{R}_+^n} \frac{\log(\prod_{i=1}^n t_i)}{(1+\sum_{i=1}^n t_i)^{n+1}}\mathrm{d}t \tag{6.70}$$

$$- n!(n+1)\int_{\mathbb{R}_+^n} \frac{\log(1+\sum_{i=1}^n t_i)}{(1+\sum_{i=1}^n t_i)^{n+1}}\mathrm{d}t =: I_1 + I_2 + I_3 \tag{6.71}$$

Using what we computed for the normalization constant we immediately get $I_1 = \log(n!)$. For I_2 we can use the symmetry of the integral to get

$$I_2 = n!n\int_{\mathbb{R}_+^n} \frac{\log(t_1)}{(1+\sum_{i=1}^n t_i)^{n+1}}\mathrm{d}t. \tag{6.72}$$

For every variable except t_1 we can follow the same procedure as for the normalization constant and get

$$I_2 = n\int_0^\infty \frac{\log(t_1)}{(1+t_1)^2}\mathrm{d}t_1 = n\left[-\frac{\log(t_1)}{1+t_1}\right]_{t_1=0}^\infty + n\int_0^\infty \frac{1}{t_1(1+t_1)}\mathrm{d}t \tag{6.73}$$

$$= n\log(0) + n\int_0^\infty -\frac{1}{t_1} + \frac{1}{1+t_1}\mathrm{d}t = n\log(0) - n\log(0) = 0 \tag{6.74}$$

where the steps involving infinities are formal and appropriate limits are suppressed for brevity. Moving on to I_3 we have

$$I_3 = -n!(n+1)\int_{\mathbb{R}_+^n} \frac{\log(1+\sum_{i=1}^n t_i)}{(1+\sum_{i=1}^n t_i)^{n+1}}\mathrm{d}t, \tag{6.75}$$

once again performing the substitution

$$\sum_{i=1}^n t_i = u_1 \tag{6.76}$$

$$t_k = u_k \text{ for } k > 1 \tag{6.77}$$

$$\mathrm{d}t = \mathrm{d}u \tag{6.78}$$

we get

$$I_3 = -n!(n+1)\int_{\mathbb{R}_+^{n-1}}\int_{\sum_{k=2}^n}^\infty \frac{\log(1+u_1)}{(1+u_1)^{n+1}}\mathrm{d}t. \tag{6.79}$$

Using

$$\int \frac{\log(1+x)}{(1+x)^{n+1}} = -\frac{1}{n}\frac{\log(1+x)}{(1+x)^n} + \frac{1}{n}\int \frac{1}{(1+x)^{n+1}} = -\frac{1}{n}\frac{\log(1+x)}{(1+x)^n} - \frac{1}{n^2}\frac{1}{(1+x)^n} \tag{6.80}$$

we get

$$I_3 = -n!(n+1)\left[\frac{1}{n}\int_{\mathbb{R}_+^{n-1}} \frac{\log(1+\sum_{i=2}^n t_i)}{(1+\sum_{i=2}^n t_i)^n}\mathrm{d}t + \frac{1}{n^2}\int_{\mathbb{R}_+^{n-1}} \frac{1}{(1+\sum_{i=2}^n t_i)^n}\mathrm{d}t\right] \tag{6.81}$$
$$= -n!(n+1)\left[\frac{1}{n}\int_{\mathbb{R}_+^{n-1}} \frac{\log(1+\sum_{i=2}^n t_i)}{(1+\sum_{i=2}^n t_i)^n}\mathrm{d}t + \frac{1}{n^2(n-1)!}\right]. \tag{6.82}$$

The integral which we are left with is just the one we started with but with n reduced to $n-1$. Repeating the above steps gives

$$I_3 = -n!(n+1)\left[\frac{1}{n!}\left(\frac{\log(1+u)}{1+u} - \frac{1}{1+u}\right)_{t=0}^{\infty} + \sum_{k=2}^{n} \frac{k!}{n!k(k-1)!}\right] \tag{6.83}$$
$$= -(n+1)\sum_{k=1}^{n}\frac{1}{k}. \tag{6.84}$$

Putting it all together we get

$$D(\mu) = -(n+1)\sum_{k=1}^{n}\frac{1}{k} + \log(n!). \tag{6.85}$$

Worth noting is that this expression grows asymptotically linear with coefficient

$$D(\mu)/n = -\frac{n+1}{n}\sum_{k=1}^{n}\frac{1}{k} + \frac{1}{n}\log(n!) \tag{6.86}$$
$$= -\frac{n+1}{n}\sum_{k=1}^{n}\frac{1}{k} + \frac{1}{n}(n\log(n) - n + O(\log(n))) \tag{6.87}$$
$$= -\sum_{k=1}^{n}\frac{1}{k} + \log(n) - 1 + O(\log(n)/n) \xrightarrow[n\to\infty]{} -1-\gamma \tag{6.88}$$

where we have used Stirling's approximation and γ is the Euler-Mascheroni constant.

This can be compared to the entropy the other type of Fano manifold we know in every dimension, namely $(\mathbb{C}\mathrm{P}^1)^n$. It has entropy

$$D(\mu_{(\mathbb{C}\mathrm{P}^1)^n}) = -2n \tag{6.89}$$

due to the additive nature of the entropy of product measures. By a conjecture of the supervisor, the toric Fano manifold minimizing the free energy times the volume of the fundamental polytope (which is in fact nothing other than the volume of the complex manifold) is minimized in every dimension for $\mathbb{C}\mathrm{P}^n$. Using the (conditional) Theorem 2 we get for $\mathbb{C}\mathrm{P}^n$ at least asymptotically

$$\mathcal{F}_{\mathbb{C}\mathrm{P}^n} = \mathrm{Vol}(\mathbb{C}\mathrm{P}^n)F(\mu_{\mathbb{C}\mathrm{P}^n}) = \frac{(n+1)^n}{n!}\left(n - (n+1)\sum_{k=1}^{n}\frac{1}{k} + \log(n!)\right) \sim \frac{-\gamma n e^n}{\sqrt{2\pi n}}. \tag{6.90}$$

Using another version of Stirling's approximation while for $(\mathbb{C}\mathrm{P})^n$

$$\mathcal{F}_{(\mathbb{C}\mathrm{P}^1)^n} = \mathrm{Vol}((\mathbb{C}\mathrm{P}^1)^n)F(\mu_{(\mathbb{C}\mathrm{P}^1)^n}) = 2^n(n-2n) \sim -n2^n \tag{6.91}$$

so the conjecture holds in this particular case at least asymptotically. There also seems to be strong numerical support for it to hold for all n in this case.

Bibliography

[1] T. Aubin, “Equations du type monge-ampère sur les varietes kähleriennes compactes,” *C. R. Acad. Sci. Paris*, no. A-B. 283 (3): Aiii, A119–A121, 1976.

[2] S.-T. Yau, “Calabi’s conjecture and some new results in algebraic geometry,” *Proceedings of the National Academy of Sciences*, vol. 74, no. 5, pp. 1798–1799, 1977, ISSN: 0027-8424. DOI: 10.1073/pnas.74.5.1798. eprint: https://www.pnas.org/content/74/5/1798.full.pdf. [Online]. Available: https://www.pnas.org/content/74/5/1798.

[3] S.-T. Yau, “On the ricci curvature of a compact kähler manifold and the complex monge-ampére equation, i,” *Communications on Pure and Applied Mathematics*, vol. 31, no. 3, pp. 339–411, 1978. DOI: https://doi.org/10.1002/cpa.3160310304. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.3160310304. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160310304.

[4] X.-X. Chen, S. Donaldson, and S. Sun, *Kahler-einstein metrics and stability*, 2012. arXiv: 1210.7494 [math.DG].

[5] R. J. Berman, *K-polystability of q-fano varieties admitting kahler-einstein metrics*, 2015. arXiv: 1205.6214 [math.DG].

[6] R. J. Berman and B. Berndtsson, *Real monge-ampere equations and kahler-ricci solitons on toric log fano varieties*, 2012. arXiv: 1207.6128 [math.DG].

[7] R. J. Berman, S. Boucksom, P. Eyssidieux, V. Guedj, and A. Zeriahi, *Kähler-einstein metrics and the kähler-ricci flow on log fano varieties*, 2016. arXiv: 1111.7158 [math.CV].

[8] M. Nakahara, *Geometry, topology and physics*, ser. Graduate student series in physics. Bristol: Hilger, 1990. [Online]. Available: https://cds.cern.ch/record/206619.

[9] P. Griffiths and J. Harris, *Principles of algebraic geometry.* New York: Wiley, 1978.

[10] T. Rockafellar, *Convex analysis.* Princeton, New Jersey: Princeton university press, 1970.

[11] P. M. Lee, “Convergence of probability measures. by patrick billingsley. pp. xii, 253. 117s. 1968. (john wiley &; sons, inc.),” *The Mathematical Gazette*, vol. 54, no. 388, p. 59, 1970. DOI: 10.2307/3612158.

[12] C. Villani, “Optimal transport – old and new,” in. 2008-01, vol. 338, p. 53. DOI: 10.1007/978-3-540-71050-9.

[13] F. Santambrogio, *Dealing with moment measures via entropy and optimal transport*, 2015. arXiv: 1507.04187 [math.FA].