AN INVERSE PROBLEM FOR THE REFRACTIVE SURFACES WITH COAXIAL LIGHTING

ARAM L. KARAKHANYAN

Abstract. In this article we examine the regularity of two types of weak solutions to a Monge-Ampère type equation which emerges in a problem of finding surfaces that refract coaxial light rays emitted from source domain and striking a given target set after refraction. Historically, ellipsoids and hyperboloids of revolution were the first surfaces to be considered in this context. The mathematical formulation commences with deriving the energy conservation equation for sufficiently smooth surfaces, regarded as graphs of functions to be sought, and then studying the existence and regularity of two classes of suitable weak solutions constructed from envelopes of hyperboloids or ellipsoids of revolution. Our main result in this article states that under suitable conditions on source and target domains and respective intensities these weak solutions are smooth.

Contents

1. Introduction

Let $\mathcal{U} \subset \mathbb{R}^n$ be a bounded domain with smooth boundary and $u : \mathcal{U} \to \mathbb{R}$ a smooth function. By Γ_u we denote the graph of u. Let γ denote the unit normal of Γ_u. We think of Γ_u as a surface that dissevers two distinct media.

From each $x \in \mathcal{U}$ we issue a ray ℓ_x parallel to e_{n+1}−the unit direction of the x_{n+1} axis in $\mathbb{R}^{n+1}$. Then ℓ_x strikes Γ_u, the surface separating the two media I and II, refracts into the second media II and strikes the receiver surface Σ, see Figure 1. Let Y be the unit direction of the refracted ray.

2010 *Mathematics Subject Classification.* Primary 35J96, Secondary 78A05, 78A46, 78A50.

Keywords: Monge-Ampère type equations, parallel refractor, antenna design.

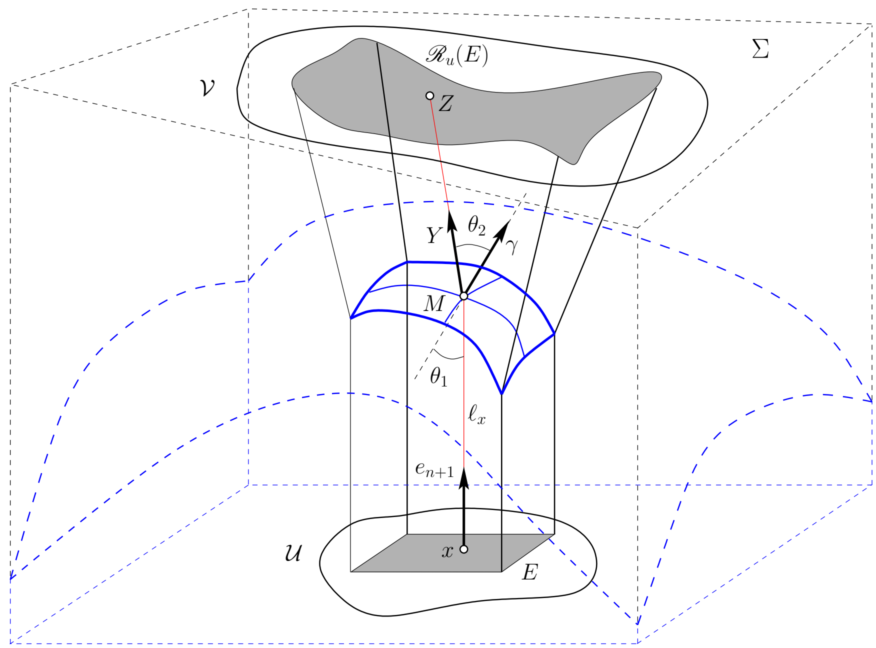


FIGURE 1. The blue doted lines confine the boundary of media I.

If γ is the unit normal at $M = (x, u(x)) \in \mathbb{R}^{n+1}$ where ℓ_x strikes Γ_u then from the refraction law we have

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1}, \tag{1.1}$$

where n_1, n_2 are the refractive indices of the media I and II respectively, dissevered by the interface Γ_u, θ_1 and θ_2 are the angles between ℓ_x and γ, and between Y and γ, respectively, see Figure 1.

Suppose that the intensity of light on $\mathcal{U}$ is $f \geq 0$ and let $\mathcal{V}$ be the set of points where the refracted rays strike the receiver Σ. Denote by $g \geq 0$ the gain intensity on $\mathcal{V}$. For each $\mathcal{U}' \subset \mathcal{U}$ let $\mathcal{V}'$ be the set of points where the rays, issued from $\mathcal{U}'$ and refracted off Γ_u, strike Σ. Thus u generates the refractor mapping

$$Z_u : \mathcal{U} \longrightarrow \mathcal{V}$$

and the illuminated domain on Σ corresponding to $\mathcal{U}' \subset \mathcal{U}$ is $\mathcal{V}' = Z_u(\mathcal{U}')$. If Γ_u is a perfect refractor, then one would have the energy balance equation (in local form)

$$\int\limits_{\mathcal{U}'} f = \int\limits_{\mathcal{V}'=Z_u(\mathcal{U}')} g. \tag{1.2}$$

The main problem that we are concerned with is formulated below:

Problem. *Assume that we are given a smooth surface Σ in $\mathbb{R}^{n+1}$, a pair of bounded smooth domains $\mathcal{U} \subset \Pi = \{X \in \mathbb{R}^{n+1} : X^{n+1} = 0\}$ and $\mathcal{V} \subset \Sigma$ and a pair of nonnegative, integrable functions $f : \mathcal{U} \to \mathbb{R}$ and $g : \mathcal{V} \to \mathbb{R}$ such that the energy balance condition holds*

$$\int_{\mathcal{U}} f = \int_{\mathcal{V}} g d\mathcal{H}^n. \tag{1.3}$$

Find a function $u : \mathcal{U} \to \mathbb{R}$ such that the following two conditions are fulfilled

$$\text{(RP)} \qquad \begin{cases} \int\limits_{\mathcal{U}'} f = \int\limits_{Z_u(\mathcal{U}')} g, \ \textit{for any measurable } \mathcal{U}' \subset \mathcal{U} \\ Z_u(\mathcal{U}) = \mathcal{V}. \end{cases}$$

Problems of this kind appear in geometric optics [11] page 315. In the 17th century Descartes posed a similar problem with target set $\mathcal{V}$ being a single point, say $\mathcal{V} = \{Z_0\}$. It was observed that the ellipsoids and hyperboloids of revolution with focal axis parallel to e_{n+1} will solve this problem if Z_0 is one of the foci. The case of general target $\mathcal{V}$ can be treated via approximation argument, namely by constructing a solution from ellipsoids or hyperboloids for finite set $\mathcal{V} = \{Z_1, \dots, Z_m\}$ and then letting $m \to \infty$. Moreover, the eccentricity of these surfaces is fixed and determined by the refractive indices n_1 and n_2. To see this we take advantage of some well-known facts from geometric optics and record them here for further reference, see [15]. Let $H(x) = Z^{n+1} - a\varepsilon - \frac{a}{b}\sqrt{b^2 + |x - x_0|^2}$ be the lower sheet of a hyperboloid of revolution with focal axis passing through the point $x_0 \in \mathcal{U}$ and parallel to e_{n+1}, see Section 7. Similarly, we define the lower half of an ellipsoid of revolution $E(x) = Z^{n+1} - a\varepsilon - \frac{a}{b}\sqrt{b^2 - |x - x_0|^2}$. If n_1 and n_2 are the refractive indices of media I and II respectively then

$$\varepsilon = \frac{n_1}{n_2} = \frac{\sin\theta_2}{\sin\theta_1} = \begin{cases} \frac{\sqrt{a^2 - b^2}}{a} < 1 & \text{for ellipsoids,} \\ \frac{\sqrt{a^2 + b^2}}{a} > 1 & \text{for hyperboloids.} \end{cases} \tag{1.4}$$

Here ε is the eccentricity, see [15]. Since ε is fixed we can drop the dependence of E and H from $b = a\sqrt{|\varepsilon^2 - 1|}$ and take

$$E(x, a, Z) = Z^{n+1} - a\varepsilon - a\sqrt{1 - \frac{(x - z)^2}{a^2(1 - \varepsilon^2)}}, \qquad \text{if } \varepsilon < 1, \tag{1.5}$$

$$H(x, a, Z) = Z^{n+1} - a\varepsilon - a\sqrt{1 + \frac{(x - z)^2}{a^2(\varepsilon^2 - 1)}}, \qquad \text{if } \varepsilon > 1. \tag{1.6}$$

We also define the constant

$$\kappa = \frac{\varepsilon^2 - 1}{\varepsilon^2} \tag{1.7}$$

which will prove to be useful, in a number of computation to follow.

Let Σ be the receiver surface given implicitly

$$\Sigma = \{Z \in \mathbb{R}^{n+1} : \psi(Z) = 0\} \tag{1.8}$$

where $\psi : \mathbb{R}^{n+1} \to \mathbb{R}$ is a smooth function. If $u \in C^2(\mathcal{U})$ then the first condition in (**RP**), after using change of variables, results a Monge-Ampère type equation for u, whereas the second one plays the role of boundary condition for u. More precisely we have the following

Theorem A. *Let $u \in C^2(\mathcal{U})$ be a solution to (**RP**). Then*

1° $Y = \varepsilon\left(\frac{\kappa Du}{1+q}, 1 - \frac{\kappa}{1+q}\right)$ *is the init direction of refracted ray,*

2° *u solves the equation*

$$\left| det \left[\frac{q+1}{t\varepsilon\kappa} \left\{ \mathrm{Id} - \kappa\varepsilon^2 Du \otimes Du \right\} + D^2 u \right] \right| = \left| -\varepsilon q \left[\frac{q+1}{t\varepsilon\kappa} \right]^n \frac{\nabla\psi \cdot Y}{|\nabla\psi|} \frac{f}{g} \right|, \tag{1.9}$$

where

$$q(x) = \sqrt{1 - \kappa(1 + |Du|^2)}, \quad \kappa = \frac{\varepsilon^2 - 1}{\varepsilon^2} \tag{1.10}$$

and t is the stretch function defined in (3.10) via an implicit relation $\psi(x + e_{n+1}u(x) + Yt) = 0$.

If the receiver Σ is a plane then taking $\psi(Z) = Z\cdot\xi + \xi_1$ we find that $t = -[Y\cdot\xi_0]^{-1}(x + u(x)e_{n+1} + \xi_1)$. In particular for the horizontal plane $X^{n+1} = m$, with some constant $m > 0$, one has

$$t = \frac{m-u}{Y^{n+1}} = (m-u)\frac{q+1}{\varepsilon(1-\kappa+q)}.$$

Quadric Σ is another example of receiver for which t can be computed explicitly. In general t is a function of $x, u(x)$ and $Du(x)$ which may not have simple explicit form. However, in terms of applications the case of planar receiver is of particular interest, since the flat screens are easy to construct. The method of the stretch function was introduced in [9, 10] to treat the near-field reflection problem. The equation for a near-field *refraction* refraction problem with point source is derived in [5], [7].

Next, we need to introduce the notion of weak solution of (1.9). It will allow us to develop the existence theory along the lines of the classical Monge-Ampère equation. To this end, we say that $u : \mathcal{U} \to \mathbb{R}$ is upper (resp. lower) admissible with respect to $\mathcal{V}$ if for any $x \in \mathcal{U}$ there is a hyperboloid $H(\cdot, a, Z)$ (resp. ellipsoid $E(\cdot, a, Z)$) with focus $Z \in \mathcal{V}$ such that $H(\cdot, a, Z)$ (resp. $E(\cdot, a, Z)$) touches u from above (resp. below) at x. Such $H(\cdot, a, Z)$ (resp. $E(\cdot, a, Z)$) is called supporting hyperboloid (resp. ellipsoid) of u at x. To fix the ideas we consider the class of upper admissible function and denote it by $\overline{\mathbb{W}}_{\rm H}(\mathcal{U}, \mathcal{V})$. The class of lower admissible functions is denoted by $\underline{\mathbb{W}}_{\rm E}(\mathcal{U}, \mathcal{V})$. For each $u \in \overline{\mathbb{W}}_{\rm H}(\mathcal{U}, \mathcal{V})$ we define the mapping $\mathscr{S}_u : \mathcal{V} \to \mathcal{U}$ by

$$\mathscr{S}_u(Z) = \{x \in \mathcal{U} : \exists a > 0 \text{ such that } H(\cdot, a, Z) \text{ is a supporting hyperboloid of } u \text{ at } x\},$$

and take

$$\beta_{u,f}(E) = \int_{\mathscr{S}_u(E)} f(x)dx, \quad E \subset \mathcal{V}.$$

Furthermore, we also consider the mapping $\mathscr{R}_u : \mathcal{U} \to \mathcal{V}$ defined by

$$\mathscr{R}_u(x) = \{Z \in \mathcal{V} : \text{there is a supporting hyperboloid } H(\cdot, a, Z) \text{ of } u \text{ at } x\}$$

and associate the following set function

$$\alpha_{u,g}(E) = \int_{\mathscr{R}_u(E)} g d\mathcal{H}^n, \quad E \subset \mathcal{V}.$$

Notice that for smooth u, the mapping $\mathscr{S}_u$ is the inverse of $\mathscr{R}_u$.

With the aid of these set functions $\alpha_{u,g}$ and $\beta_{u,f}$ we can introduce two notions of weak solution to (**RP**), called A and B type weak solutions, respectively. It is not hard to see that $\beta_{u,f}$ is in fact σ-additive measure, while for $\alpha_{u,g}$ it is less obvious. Towards proving this the major obstruction is to show that $\mathscr{R}_u$ is one-to-one modulo a set of vanishing $\mathcal{H}^n$ measure on Σ. This is circumvented by introducing the Legendre-like transformation $v(z)$ of an admissible function $u(x)$ in Section 10 defined as an upper envelope of some function of $\mathrm{dist}(Z, X)$ for $Z \in \mathcal{V}$ and $X \in \mathcal{U}$. In order to infer that $v(z)$ is semi-concave (which in turn will lead to σ-additivity of $\alpha_{u,g}$) we assume that (1.12) is fulfilled. That done, one can show that an A-type weak solution exists in the sense of Definition 10.2.

If, for a moment, we take the existence of A-type weak solution for granted, the question about its regularity is even more complex. To set stage for the weak solutions we assume that $\Sigma = \{Z \in \mathbb{R}^{n+1} : \psi(Z) = 0\}$ and $\psi : \mathbb{R}^{n+1} \to \mathbb{R}$ being a smooth function. Clearly, some conditions must be imposed on ψ to guarantee, among other things, that the right hand side of the equation (1.9) is well defined, at least for smooth solutions.

To this end we enlist the following conditions to be used in the construction of weak solutions and proving their smoothness.

(1.11) $\nabla\psi(Z)\cdot(X-Z)>0\quad\forall X\in(\mathcal{U}\times[0,m_0]),\forall Z\in\Sigma$ and for some large constant $m_0>0$,

(1.12) $\mathrm{dist}(\mathcal{U},\mathcal{V})>0$,

(1.13) $\mathcal{V}$ is $R-$convex with respect to $\mathcal{U}$, see Definition 9.2,

(1.14) $f,g>0$,

(1.15) $\frac{1}{t}\left[\frac{t\varepsilon\kappa}{q+1}\right]^2\mathrm{II}+\frac{\kappa}{q}\frac{\psi_{n+1}}{|\nabla\psi|}\left(\mathrm{Id}+\kappa\frac{p\otimes p}{q^2}\right)>0,\quad$ if $\kappa>0$,

where II is the second fundamental form of Σ. The subdomain of $\mathcal{U}\times[0,\infty)$ where (1.11)-(1.15) are simultaneously satisfied is called the *regularity domain* $\mathcal{D}$.

It is worthwhile to explain the meaning of these conditions: the first one (1.11) means that the reflected rays are not striking Σ tangentially, otherwise Σ would not detect the gain intensity at the tangential points, i.e. at the points where $\nabla\psi(Z)\cdot(X-Z)=0$. On the technical level, however, it allows to apply the inverse function theorem to recover the stretch function $t=t(x,u,Du)$. It is worth pointing out that (1.11) holds for a large class of surfaces Σ. To see this we first notice that there is a positive constant $c(\varepsilon)$, depending only on ε such that $Y^{n+1}\in(c(\varepsilon),1]$. In other words the unit directions Y of refracted rays are within the cone $c(\varepsilon)<Y^{n+1}\le 1$. Indeed, $Y^{n+1}=\varepsilon[1+(\sqrt{\cos^2\theta_1-\kappa}-\cos\theta_1)\gamma^{n+1}]$ from refraction law, see (3.4) and Figure 1. If u is not differentiable at x, we interpret γ as one of the normals of supporting planes of admissible u at $x\in\mathcal{U}$ since u is concave (resp. convex) if u is upper (resp. lower) admissible. Consequently if u is lower admissible then $Y^{n+1}\ge\varepsilon$ if $\kappa<0$, i.e. $\varepsilon<1$ and hence $c(\varepsilon)=\varepsilon<1$. On the other hand if $\kappa>0$ then for any $u\in\mathbb{H}(\mathcal{U},\mathcal{V})$ we have

$$|Du|<\frac{1}{\sqrt{\varepsilon^2-1}}\quad\text{if }\kappa=\frac{\varepsilon^2-1}{\varepsilon^2}>0. \tag{1.16}$$

This simply follows from the fact that supporting hyperboloids control the magnitude of the gradient of u. But in its turn $|DH|$ of any hyperboloid H given by (1.6) satisfies the estimate (1.16). Because $\gamma^{n+1}=\cos\theta_1$ (see Figure 1 and the derivation of (3.7)) we infer that

$$(\sqrt{\cos^2\theta_1-\kappa}-\cos\theta_1)\gamma^{n+1}=\frac{-\kappa}{\sqrt{\cos^2\theta_1-\kappa}+\cos\theta_1}\gamma^{n+1}>-\kappa$$

and consequently $Y^{n+1}>\varepsilon(1-\kappa)=\varepsilon^{-1}<1$. Thus for $\varepsilon>1$ we can take $c(\varepsilon)=\varepsilon^{-1}$. From here we see that (1.11) holds for any horizontal receiver $Z^{n+1}=m$, for large $m>0$. More generally if Σ is concave in Z^{n+1} direction and the normal mapping of Σ is strictly inside of the cone $c(\varepsilon)<Y^{n+1}$ on the unit sphere then (1.11) holds true. This leads to the following cone condition for the unit directions of refracted rays

$$0<c(\varepsilon)\le Y^{n+1}\le 1 \tag{1.17}$$

The second condition (1.12) assures that the Legendre-like transformation $v(z)$ for an admissible function u is well defined as an envelope of C^1 smooth functions, in particular $u^\star$ is semi-concave and hence differentiable almost everywhere, see Section 10. This yields that $\alpha_{u,g}$ is a Radon measure.

The next two conditions (1.13) and (1.14) assure that B-type solution is also of A-type and therefore one gets the existence of A-type weak solutions in some indirect way using the methods of [4], [20]. That done, we can approximate $\mathcal{V}$ by R-convex domains and show the existence of A-type weak solutions without assuming (1.13), see Theorem C4.

Last condition (1.15), which is crucial for regularity of weak solutions, deserves special attention because is the most sophisticated one and in order to verify it we have formulated the following

Theorem B. *Let u be a $C^2(\mathcal{U})$ solution of (1.9) and $||$ be the second fundamental form of $\Sigma = \{Z \in \mathbb{R}^{n+1} : Z^{n+1} = \varphi(z)\}$. If $\kappa > 0, \mathrm{II} \geq 0$ of $\kappa < 0, \mathrm{II} \leq 0$ then (1.15) holds true with $c_0 > 0$ depending only on ε and the Lipschitz norm of u.*

If Σ is a graph, say $Z^{n+1} = \varphi(z)$ then (1.15) can be rewritten as

$$\frac{1}{t}\left[\frac{t\varepsilon\kappa}{q+1}\right]^2\sqrt{1+|D\varphi|^2}\mathrm{II}+\frac{\kappa}{q}\left(\mathrm{Id}+\kappa\frac{p\otimes p}{q^2}\right)>0, \quad \text{if } \kappa>0.$$

In lieu of (1.17) this assumption on Σ is restrictive. In addition, Theorem B suggest that it is convenient to think of Σ as an unbounded convex (reap. concave) surface without boundary if $\kappa > 0$ (reap. $\kappa < 0$) by extending φ to $\mathbb{R}^n$ as a convex function $\widetilde{\varphi}$ such that $\varphi(z) \to \pm\infty$ as $|z| \to \infty$. We will take advantage of such extension of φ (and hence Σ) in Section 7.4 and Lemma 9.2, see also Remark 6.1.

Now we are ready to formulate our main existence result.

Theorem C. 1 *if $f, g \geq 0$ and (1.3) holds then there is a B-type weak solution provided that the condition below*

$$Z^{n+1} \geq \left[\frac{2}{\varepsilon-1}+\frac{1}{\sqrt{\varepsilon^2-1}}\right]\rho(z) \tag{1.18}$$

is satisfied. Here $\rho(z) = \inf\{R > 0 : \mathcal{U} \subset B_z(R)\}$ is the maximal visibility radius from $z \stackrel{def}{=} \widehat{Z} \in \widehat{\mathcal{V}}$,

2 *if (1.11) and (1.12) hold then $\alpha_{u,g}$ is countably additive,*

3 *if (1.11)-(1.13), (1.18) hold and $f \geq 0$ while $g > 0$ then B-type weak solution is also of A type,*

4 *if we remove the $R-$convexity assumption but require the positivity of densities (1.14) and (1.11)-(1.12), (1.18) then again any B-type weak solution is also of A-type.*

The proof of Theorem C1 is by polyhedral approximation while utilising the confocal expansion of hyperboloids as described in Section 7.4. In this regard the condition (1.18) in Theorem C1 says that one can construct a B-type weak solution if there is sufficient span between Π and Σ.

Our last result concerns with the smoothness of A-type weak solutions. We use the method of comparing mollified weak solution with the solution of Dirichlet problem to the regularised equation in a small ball B. To this end one first has to obtain $C^{2,\alpha}, \alpha \in (0, 1]$ estimates in $\overline{B}$ for the solutions of mollified equations and second to have uniform C^2 estimates in, say, $\frac{1}{2}B$. Then passing to limit and using the comparison principle the result will follows. The construction of weak solutions to Dirichlet's problem is based on Perron's method and follows along the lines developed in the paper by Xu-Jia Wang [21] where a reflector design problem is studied. Our research is inspired by [21] and subsequent developments in [9], [10] [8]. For more recent results on this problem see [12]. The global C^2 estimates for the solution of Dirichlet's problem for the regularised equation follow from [6] whereas the local uniform estimates in $\frac{1}{2}B$ are established in [14]. Thus we have the following theorem

Theorem D. *Let f, g be C^2 smooth functions such that $\lambda \leq f, g \leq \Lambda$ for some constants $\Lambda > \lambda > 0$ and the conditions (1.11)- (1.15) are satisfied. Then A-type weak solutions of (**RP**) are locally C^2 regular in $\mathcal{U}$.*

The conditions (1.11)- (1.15) cannot be relaxed as one may easily construct counterexamples to regularity in the spirit of those in [9], [10]. For instance let us examine (1.13) (see also Remark 11.3), if we take the two point target $\mathcal{V} = \{Z_1\} \cup \{Z_2\}$ and consider $H(x) = \min[H(x, a_1, Z_1), H(x, a_2, Z_2)]$ such that these hyperboloids have non empty intersection over $\mathcal{U}$, then approximating $\mathcal{V}$ by smooth R-convex sets $\mathcal{V}_t$ we obtain a sequence of admissible H_t converging to H as $t \to 0$. But if t is sufficiently close to 0 then H_t cannot remain C^1 smooth because otherwise the limit H would also be C^1 which is impossible., see [9] for more details.

The rest of the paper is organized as follows: in the next section we derive the main formulae. Then we prove Theorem A in Section 3. The main result there is Proposition 4.1 from which the proof of Theorem A easily follows, see Section 4.2. Section 5 contains some preliminary discussion on the condition (1.15) and after that in Section 6 we give the proof of Theorem B. The admissible functions are introduced in Section 7 where we also exhibit some interesting properties of hyperboloids of revolution, notably the dual admissibility and confocal expansion. Employing the polyhedral approximation technique and weak convergence of measures $\beta_{u,f}$ we prove Theorem C1 in Section 8. The first direct application of (1.15) is given in Lemma 9.1, which is G. Loeper's geometric interpretation of the MTW condition from [14]. A direct consequence of this is Lemma 9.2 on approximation of admissible function by smooth subsolutions of (1.9). This is a crucial ingredient of the proof of Theorem D. Next we introduce the Legendre-like transformation of an admissible u and conclude Theorem C2. The proofs of Theorem C3-4 follow from a comparison of A and B type weak solutions by extending the results of Luis Caffarelli [4] and John Urbas [20] for the classical Monge-Ampère equation to (1.9). This is done in Section 11. The last two sections are concerned with the higher regularity A-type weak solutions of (**RP**). Our approach is classical and closely affiliates with the classical Monge-Ampère equation for which A. Pogorelov was the first to propose it, see [16], [17]. Consequently, we first prove the solvability of weak Dirichlet's problem when the boundary data is given as the trace of an A-type weak subsolution. Uniqueness follows from comparison principle stated in Proposition 12.1. Finally in Section 13 we give the proof of our main regularity result, Theorem D.

2. Notations

$C, C_0, C_n, \cdots$	generic constants,
Π	$\Pi = \mathbb{R}^n \times \{0\}$,
$\overline{\mathcal{U}}$	closure of a set $\mathcal{U}$,
$\partial\mathcal{U}$	boundary of a set $\mathcal{U}$,
$\widehat{\mathcal{U}}$	the projection of $\mathcal{U} \subset \mathbb{R}^{n+1}$ on Π,
$\widehat{X}$	$(x_1, x_2, \dots, x_n, 0)$ projection of $X = (x_1, x_2, \dots, x_n, x_{n+1})$,
ε	eccentricity,
κ	$\kappa = \frac{\varepsilon^2 - 1}{\varepsilon^2}$,
$\mathcal{H}^n$	n dimensional Hausdorff measure on Σ,
∂_i	partial derivate with respect to x_i variable,
Du	the gradient of a function u,
$\rho(z)$	$\inf\{R > 0 : \mathcal{U} \subset B_z(R)\}$ is the maximal visibility radius from $z \stackrel{def}{=} \widehat{Z} \in \widehat{\mathcal{V}}$,
q	see (1.10),
$\mathbb{H}(\mathcal{U}, \Sigma)$	the class of hyperboloids of revolution with focus on Σ,
$\mathbb{H}^+(\mathcal{U}, \mathcal{V})$	hyperboloids from $\mathbb{H}(\mathcal{U}, \mathcal{V})$ which are nonnegative in $\mathcal{U}$,
$\overline{\mathbb{W}}_{\mathrm{H}}, \underline{\mathbb{W}}_{\mathrm{E}}$	upper and lower admissible functions, see Lemma 7.1,
$\overline{\mathbb{W}}^0_{\mathrm{H}}(\mathcal{U}, \mathcal{V})$	polyhedral admissible functions.

3. Main formulae

In this section we derive the Monge-Ampère type equation (1.9) manifesting the energy balance condition (1.2) in the refractor problem (**RP**), see Introduction.

3.1. **Computing Y.** We first compute the unit direction of the refracted ray. Denote by γ the unit normal to the graph of u, that is

$$\gamma = \frac{(-D_1 u, \dots, -D_n u, 1)}{\sqrt{1+|Du|^2}}. \tag{3.1}$$

Since ℓ_x, Y and γ lie in the same hyperplane we have

$$Y = \mathcal{A} e_{n+1} + \mathcal{B}\gamma, \tag{3.2}$$

for some coefficients $\mathcal{A}$ and $\mathcal{B}$. Computing the scalar products $Y \cdot \gamma$ and $Y \cdot e_{n+1}$ we obtain the following equations (cf. (1.1))

$$\begin{cases} \cos\theta_2 = \mathcal{A}\cos\theta_1 + \mathcal{B}, \\ \cos(\theta_1 - \theta_2) = \mathcal{A} + \mathcal{B}\cos\theta_1. \end{cases}$$

Multiplying the first equation by $\cos\theta_1$ and subtracting from the second one we conclude

$$\mathcal{A} = \frac{\sin\theta_2}{\sin\theta_1}, \qquad \mathcal{B} = \cos\theta_2 - \mathcal{A}\cos\theta_1.$$

Recalling our notations

$$\kappa = \frac{\varepsilon^2 - 1}{\varepsilon^2}, \qquad \varepsilon = \frac{n_1}{n_2}, \tag{3.3}$$

we see that $\mathcal{A} = \varepsilon$. Furthermore

$$n_2^2 - n_2^2\cos^2\theta_2 = n_2^2\sin^2\theta_2 = n_1^2\sin^2\theta_1 = n_1^2 - n_1^2\cos^2\theta_1.$$

Dividing both sides of this identity by n_2^2 we obtain

$$\cos^2\theta_2 = \varepsilon^2\cos^2\theta_1 - (\varepsilon^2 - 1) = \varepsilon^2(\cos^2\theta_1 - \kappa).$$

Therefore from $\mathcal{A} = \varepsilon$ we conclude that $\mathcal{B} = \varepsilon(\sqrt{\cos^2\theta_1 - \kappa} - \cos\theta_1)$. Returning to (3.2) we infer that the unit direction of the refracted ray is

$$Y = \varepsilon\left(e_{n+1} + (\sqrt{\cos^2\theta_1 - \kappa} - \cos\theta_1)\gamma\right). \tag{3.4}$$

Notice that (3.1) implies

$$\cos\theta_1 = \gamma \cdot e_{n+1} = \frac{1}{\sqrt{1+|Du|^2}}.$$

Consequently, denoting $Y = (Y^1, Y^2, \dots, Y^n, Y^{n+1})$ and $y \in \mathbb{R}^n$, the projection of Y onto $\Pi = \{X \in \mathbb{R}^{n+1} : X^{n+1} = 0\}$, (i.e. $y = (Y^1, Y^2, \dots, Y^n, 0)$) we get

$$\begin{aligned} y &= -\varepsilon\frac{Du}{\sqrt{1+|Du|^2}}\left(\sqrt{\cos^2\theta_1 - \kappa} - \cos\theta_1\right) \\ &= \frac{\varepsilon\kappa Du}{\sqrt{1+|Du|^2}}\frac{1}{\sqrt{\cos^2\theta_1 - \kappa} + \cos\theta_1} \\ &= \varepsilon\kappa\frac{Du}{\sqrt{1-\kappa(1+|Du|^2)}+1}. \end{aligned} \tag{3.5}$$

From this computation it follows that

$$Y^{n+1} = \varepsilon\left(1 - \frac{\kappa}{1+\sqrt{1-\kappa(1+|Du|^2)}}\right). \tag{3.6}$$

Combining (3.5) and (3.6) we obtain

$$Y = \varepsilon\left(\frac{\kappa Du}{1+\sqrt{1-\kappa(1+|Du|^2)}}, 1 - \frac{\kappa}{1+\sqrt{1-\kappa(1+|Du|^2)}}\right). \tag{3.7}$$

If we use the notation $q(x) = \sqrt{1-\kappa(1+|Du|^2)}$ (see (1.10)) then (3.7) takes the form

$$Y = \varepsilon\left(\frac{\kappa Du}{1+q}, 1-\frac{\kappa}{1+q}\right). \tag{3.8}$$

Notice that by (3.7) $Y^{n+1} > 0$ for all values of κ.

3.2. **Stretch function.** Assume that ψ is a smooth function $\psi : \mathbb{R}^{n+1} \to \mathbb{R}$, and the receiver Σ is given as the zero set of ψ

$$\Sigma = \{Z \in \mathbb{R}^{n+1} : \psi(Z) = 0\}. \tag{3.9}$$

Let us represent the mapping $Z : \mathcal{U} \to \Sigma$ in the following form

$$Z = x + e_{n+1}u(x) + Yt, \tag{3.10}$$

where $t = t(x, u(x), Du(x))$ is determined from the equation $\psi(Z) = 0$ and is called the *stretch function*. It is worthwhile to point out that the stretch function t can be explicitly computed for a wide class of elementary surfaces. For instance, if Σ is the horizontal plane $Z^{n+1} = m > 0$ then from simple geometric considerations one finds that

$$t = \frac{m-u}{Y^{n+1}}$$

where Y^{n+1} is given by (3.6).

In lemma to follow we denote by z the projection of Z onto Π, that is $z = x + ty$.

Lemma 3.1. *Let $dS_{\mathcal{U}}$ and $dS_{\mathcal{V}}$ be the area elements on $\mathcal{U}$ and $Z(\mathcal{U}) = \mathcal{V} \subset \Sigma$ respectively and z being the projection of Z onto $\Pi = \{Z \in \mathbb{R}^{n+1} : Z^{n+1} = 0\}$. Then we have*

$$\begin{aligned} J = \frac{dS_{\mathcal{V}}}{dS_{\mathcal{U}}} &= \begin{vmatrix} Z^1_1, & \cdots, & Z^1_n, & \nu^1 \\ \vdots & \ddots & \vdots & \vdots \\ Z^n_1, & \cdots, & Z^n_n, & \nu^n \\ Z^{n+1}_1, & \cdots, & Z^{n+1}_n, & \nu^{n+1} \end{vmatrix} \\ &= -\frac{|\nabla\psi|}{\psi_{n+1}} detDz, \end{aligned} \tag{3.11}$$

where ν is the unit normal of Σ.

Proof. The first equality in (3.11) follows from the change of variables formula. Differentiating the equality $\psi(Z) = 0$ by x_i we have that

$$\partial_i Z^{n+1} = -\frac{1}{\partial_{n+1}\psi}\sum_{k=1}^{n} \partial_i z^k \partial_{z_k}\psi.$$

Using this identity we multiply j-th row of matrix in (3.11) by $\partial_{z_j}\psi$ and subtract it from the $(n+1)$st row in order to get

$$\det\begin{vmatrix} Z_1^1, & \cdots, & Z_n^1, & \nu_1 \\ \vdots & \ddots & \vdots & \vdots \\ Z_1^n, & \cdots, & Z_n^n, & \nu_n \\ Z_1^{n+1}, & \cdots, & Z_n^{n+1}, & \nu_{n+1} \end{vmatrix} = -\frac{1}{\psi_{n+1}}\det\begin{vmatrix} Z_1^1, & \cdots, & Z_n^1, & \nu_1 \\ \vdots & \ddots & \vdots & \vdots \\ Z_1^n, & \cdots, & Z_n^n, & \nu_n \\ \sum_{k=1}^n \partial_1 z^k \partial_{z_k}\psi, & \cdots, & \sum_{k=1}^n \partial_n z^k \partial_{z_k}\psi, & -\psi_{n+1}\nu_{n+1} \end{vmatrix}$$

$$= -\frac{1}{\psi_{n+1}}\det\begin{vmatrix} Z_1^1, & \cdots, & Z_n^1, & \nu_1 \\ \vdots & \ddots & \vdots & \vdots \\ Z_1^n, & \cdots, & Z_n^n, & \nu_n \\ 0, & \cdots, & 0, & -\sum_{k=1}^{n+1}\psi_k\nu_k \end{vmatrix}.$$

Finally noting that $\nu = \frac{\nabla\psi}{|\nabla\psi|}$ the desired identity follows. □

Lemma 3.2. *Let $C \in \mathbb{R}$ and $\xi, \eta \in \mathbb{R}^n$. Consider the matrix $\mu = \mathrm{Id} + C\xi\otimes\eta = \delta_{ij} + C\xi^i\eta^j$ where $\mathrm{Id} = \delta_{ij}$ is the identity matrix. Then the inverse matrix of μ is*

$$\begin{aligned} det\mu &= 1 + C\xi\cdot\eta, \\ \mu^{-1} &= \mathrm{Id} - \frac{C\xi\otimes\eta}{1 + C(\xi\cdot\eta)}. \end{aligned}$$

Here and henceforth Id *is the identity matrix.*

Proof. Without loss of generality we assume that $\xi = e_1$ then $\det\mu = 1 + C\eta^1$. It is easy to check that $\mu(\mathrm{Id} - \frac{C\xi\otimes\eta}{1+C(\xi\cdot\eta)}) = \mathrm{Id}$. □

Finally, we derive a formula for the first order derivatives of the stretch function t. Let us differentiate the equation $\psi(Z) = 0$ with respect to x_j to get

$$\sum_{k=1}^n \psi_k(\delta_{kj} + t_j y^k + t y_j^k) + \psi_{n+1}(u_j + t_j Y^{n+1} + tY_j^{n+1}) = 0.$$

From here we find

$$t_j = -\frac{1}{\nabla\psi\cdot Y}[\psi_j + \psi_{n+1}u_j + t(\nabla\psi\cdot Y_j)]. \tag{3.12}$$

4. Proof of Theorem A

In this section we prove Theorem A. We begin with a computation for the matrix Dz, where z is the projection of Z on to Π.

Proposition 4.1. *Let $u \in C^2(\mathcal{U})$ and Z be the corresponding refractor map, then with the same notations as in Lemma 3.1 we have*

$$Dz = \mu_1\mu_2\left[\mathrm{Id} - \kappa\varepsilon^2 Du\otimes Du + \frac{t\kappa\varepsilon}{1+h}D^2u\right], \tag{4.1}$$

where

$$\mu_1 = \mathrm{Id} - \frac{y\otimes(\widehat{\nabla}\psi - y\frac{\psi_{n+1}}{Y^{n+1}})}{\nabla\psi\cdot Y}, \qquad \mu_2 = \mathrm{Id} + \kappa\frac{Du\otimes Du}{q(q+1)}, \tag{4.2}$$

$q=\sqrt{1-\kappa(1+|Du|^2)}$ *and*

$$\widehat{\nabla}\psi=(\psi_1,\dots,\psi_n,0). \tag{4.3}$$

In order to prove Proposition 4.1 we will need the following

Lemma 4.1. *Let $z(x), x\in\mathcal{U}$ be the projection of the mapping $Z(x)$ onto $\Pi=\{X\in\mathbb{R}^{n+1}: X^{n+1}=0\}$. Then*

$$Dz=\mu_1\left(\mathrm{Id}-y\otimes[y+DuY^{n+1}]+tDy\right) \tag{4.4}$$

where μ_1 is defined by (4.2).

Proof. Introduce the matrix

$$\mu_0=\delta_{ij}-y^i\frac{\psi_j+u_j\psi_{n+1}}{\nabla\psi\cdot Y}. \tag{4.5}$$

Using (3.12) and recalling $z=x+ty$ we compute

$$\begin{aligned} z^i_j &= \delta_{ij}+t_jy^i+ty^i_j \\ &= \delta_{ij}+ty^i_j-y^i\frac{1}{\nabla\psi\cdot Y}[\psi_j+\psi_{n+1}u_j+t(\nabla\psi\cdot Y_j)] \\ &= \underbrace{\delta_{ij}-y^i\frac{[\psi_j+u_j\psi_{n+1}]}{\nabla\psi\cdot Y}}_{\mu_0}+t\left[y^i_j-\frac{y^i(\nabla\psi\cdot Y_j)}{\nabla\psi\cdot Y}\right] \\ &= \mu_0+t\left[y^i_j-\frac{y^i(\nabla\psi\cdot Y_j)}{\nabla\psi\cdot Y}\right]. \end{aligned} \tag{4.6}$$

In order to deal with the remaining matrix we recall that $(Y^{n+1})^2=1-|y|^2$ and hence $Y^{n+1}_j=-\frac{yy_j}{Y^{n+1}}$. Consequently, setting $\widehat{\nabla}\psi=(\psi_1,\dots,\psi_n,0)$ (see (4.3)) we infer

$$\begin{aligned} y^i_j-\frac{y^i(\nabla\psi\cdot Y_j)}{\nabla\psi\cdot Y} &= y^i_j-\frac{y^i}{\nabla\psi\cdot Y}\left(\widehat{\nabla}\psi\cdot y_j-\psi_{n+1}\frac{y\cdot y_j}{Y^{n+1}}\right) \\ &= y^i_j-\frac{y^i}{\nabla\psi\cdot Y}\left[\left(\widehat{\nabla}\psi-\frac{\psi_{n+1}}{Y^{n+1}}y\right)y_j\right]. \end{aligned} \tag{4.7}$$

Combining (4.6) and (4.7) we obtain the following formula for Dz, written in intrinsic form

$$\begin{aligned} Dz &= \mu_0+t\left[\mathrm{Id}-\frac{y\otimes\left(\widehat{\nabla}\psi-\frac{\psi_{n+1}}{Y^{n+1}}y\right)}{\nabla\psi\cdot Y}\right]Dy \\ &= \mu_0+t\mu_1Dy \\ &= \mu_1(\mu_1^{-1}\mu_0+tDy) \end{aligned} \tag{4.8}$$

where the second equality follows from the definition of matrix μ_1, see (4.2).

Next, we compute μ_1^{-1}. From Lemma 3.2 and the identity $[Y^{n+1}]^2=1-|y|^2$ we get

$$\begin{aligned} \mu_1^{-1} &= \mathrm{Id}+y\otimes\frac{\widehat{\nabla}\psi-\frac{\psi_{n+1}}{Y^{n+1}}y}{\nabla\psi\cdot Y-\left(\widehat{\nabla}\psi\cdot y-|y|^2\frac{\psi_{n+1}}{Y^{n+1}}\right)} \\ &= \mathrm{Id}+\frac{Y^{n+1}}{\psi_{n+1}}y\otimes\left[\widehat{\nabla}\psi-\frac{\psi_{n+1}}{Y^{n+1}}y\right], \end{aligned} \tag{4.9}$$

where the last equality follows from the observation

$$\nabla\psi\cdot Y-\left(\widehat{\nabla}\psi\cdot y-|y|^2\frac{\psi_{n+1}}{Y^{n+1}}\right) = \psi_{n+1}Y^{n+1}+(1-(Y^{n+1})^2)\frac{\psi_{n+1}}{Y^{n+1}} = \frac{\psi_{n+1}}{Y^{n+1}}. \tag{4.10}$$

It is convenient to rewrite this identity in the following form

$$\left[\widehat{\nabla}\psi\cdot y-|y|^2\frac{\psi_{n+1}}{Y^{n+1}}\right]\frac{Y^{n+1}}{\psi_{n+1}}\frac{1}{\nabla\psi\cdot Y}=\frac{Y^{n+1}}{\psi_{n+1}}-\frac{1}{\nabla\psi\cdot Y}. \tag{4.11}$$

Consequently, we obtain

$$\begin{aligned}\mu_1^{-1}\mu_0 &= \left(\mathrm{Id}+\frac{Y^{n+1}}{\psi_{n+1}}y\otimes\left[\widehat{\nabla}\psi-\frac{\psi_{n+1}}{Y^{n+1}}y\right]\right)\left(\mathrm{Id}-y\otimes\frac{\widehat{\nabla}\psi+Du\psi_{n+1}}{\nabla\psi\cdot Y}\right)\\ &= \mathrm{Id}-y\otimes\frac{\widehat{\nabla}\psi+Du\psi_{n+1}}{\nabla\psi\cdot Y}+\frac{Y^{n+1}}{\psi_{n+1}}y\otimes\left[\widehat{\nabla}\psi-\frac{\psi_{n+1}}{Y^{n+1}}y\right]-\\ &\quad-\left[\widehat{\nabla}\psi\cdot y-|y|^2\frac{\psi_{n+1}}{Y^{n+1}}\right]\frac{Y^{n+1}}{\psi_{n+1}}\frac{1}{\nabla\psi\cdot Y}\left\{y\otimes\widehat{\nabla}\psi+Du\psi_{n+1}\right\}\end{aligned}$$

Applying (4.11) to the last term in this computation we get

$$\begin{aligned}\mu_1^{-1}\mu_0 &= \mathrm{Id}-y\otimes\frac{\widehat{\nabla}\psi+Du\psi_{n+1}}{\nabla\psi\cdot Y}+\frac{Y^{n+1}}{\psi_{n+1}}y\otimes\left[\widehat{\nabla}\psi-\frac{\psi_{n+1}}{Y^{n+1}}y\right]-\\ &\quad-\left[\frac{Y^{n+1}}{\psi_{n+1}}-\frac{1}{\nabla\psi\cdot Y}\right]\left\{y\otimes\widehat{\nabla}\psi+Du\psi_{n+1}\right\}\\ &= \mathrm{Id}+\frac{Y^{n+1}}{\psi_{n+1}}y\otimes\left[\widehat{\nabla}\psi-\frac{\psi_{n+1}}{Y^{n+1}}y\right]-\\ &\quad-\frac{Y^{n+1}}{\psi_{n+1}}\left\{y\otimes\widehat{\nabla}\psi+Du\psi_{n+1}\right\}\\ &= \mathrm{Id}-y\otimes[y+Du\psi_{n+1}].\end{aligned}$$

Plugging in the computed form of $\mu_1^{-1}\mu_0$ into (4.8) the result follows. □

4.1. **Proof of Proposition 4.1.** To finish the proof of Proposition 4.1, it remains to express Dz through the Hessian D^2u. We have from (3.8)

$$y = \varepsilon\kappa\frac{Du}{q+1}, \tag{4.12}$$

$$Y^{n+1} = \varepsilon\left(1-\frac{\kappa}{q+1}\right), \tag{4.13}$$

where $q=\sqrt{1-\kappa(1+|Du|^2)}$, see (1.10). From the definition of q we have $Dq=-\kappa DuD^2u/q$, thus

$$\begin{aligned}Dy &= \varepsilon\kappa\left[\mathrm{Id}+\kappa\frac{Du\otimes Du}{q(q+1)}\right]\frac{D^2u}{q+1}\\ &= \varepsilon\kappa\mu_2\frac{D^2u}{q+1},\end{aligned}$$

where μ_2 is the matrix in (4.2). Now Lemma 4.1 yields

$$
\begin{aligned}
Dz &= \mu_1\left(\mathrm{Id} - y\otimes[y+DuY^{n+1}] + t\varepsilon\kappa\mu_2\frac{D^2u}{q+1}\right) \\
&= \mu_1\mu_2\left(\mu_2^{-1}\left\{\mathrm{Id} - y\otimes[y+DuY^{n+1}]\right\} + t\varepsilon\kappa\frac{D^2u}{q+1}\right) \\
&= \mu_1\mu_2\left(\mu_2^{-1}\mathcal{M} + t\varepsilon\kappa\frac{D^2u}{q+1}\right)
\end{aligned}
\tag{4.14}
$$

where $\mathcal{M} = \mathrm{Id} - y\otimes[y+DuY^{n+1}]$.

Using (4.12) we can further simplify the matrix $\mathcal{M} = \mathrm{Id} - y\otimes[y+DuY^{n+1}]$ to get

$$
\begin{aligned}
\mathcal{M} &= \mathrm{Id} - y\otimes(y+DuY^{n+1}) \\
&= \mathrm{Id} - \frac{\varepsilon^2\kappa^2}{(1+q)^2}Du\otimes Du - \frac{\varepsilon^2\kappa(1-\frac{\kappa}{1+q})}{1+q}Du\otimes Du \\
&= \mathrm{Id} - \frac{\varepsilon^2\kappa}{1+q}Du\otimes Du.
\end{aligned}
\tag{4.15}
$$

By Lemma 3.2 we have for the inverse of μ_2 (see (4.2))

$$
\begin{aligned}
\mu_2^{-1} &= \mathrm{Id} - \frac{\kappa}{q^2+q+\kappa|Du|^2}Du\otimes Du \\
&= \mathrm{Id} - \frac{\kappa}{1-\kappa+q}Du\otimes Du,
\end{aligned}
\tag{4.16}
$$

where the last equality follows from the definition of q, see (1.10). It remains to compute $\mu_2^{-1}\mathcal{M}$. From (4.16) and (4.15) we obtain

$$
\begin{aligned}
\mu_2^{-1}\mathcal{M} &= \left[\mathrm{Id} - \frac{\kappa}{1-\kappa+q}Du\otimes Du\right]\left[\mathrm{Id} - \frac{\varepsilon^2\kappa}{1+q}Du\otimes Du\right] \\
&= \mathrm{Id} + [I + II + III]\,Du\otimes Du
\end{aligned}
$$

where

$$
\begin{aligned}
I &= -\frac{\kappa}{1-\kappa+q}, \\
II &= -\frac{\varepsilon^2\kappa}{1+q}, \\
III &= \frac{\varepsilon^2\kappa^2|Du|^2}{(1+q)(1-\kappa+q)}.
\end{aligned}
$$

It follows from (1.10) that $-\kappa|Du|^2 = q^2 - 1 + \kappa$, therefore

$$
III = \frac{\varepsilon^2\kappa(-q^2+1-\kappa)}{(1+q)(1-\kappa+q)}.
$$

Adding this to II we have

$$
\begin{aligned}
II + III &= \frac{\varepsilon^2\kappa}{1+q}\left[-1 + \frac{-q^2+1-\kappa}{1-\kappa+q}\right] \\
&= -\frac{q\varepsilon^2\kappa}{1-\kappa+q}.
\end{aligned}
$$

Finally we compute the total sum

$$\begin{aligned} I+II+III &= -\frac{\kappa}{1-\kappa+q}-\frac{q\varepsilon^2\kappa}{1-\kappa+q} \\ &= -\frac{\kappa}{1-\kappa+q}\left[q\varepsilon^2+1\right] \\ &= -\frac{\kappa}{1-\kappa+q}\left[\frac{q}{1-k}+1\right] \\ &= -\frac{\kappa}{1-\kappa} \\ &= -\kappa\varepsilon^2, \end{aligned}$$

where the last line follows from the definition of κ, see (3.3).

Returning to (4.14) and utilising these computations we get

$$\begin{aligned} Dz &= \mu_1\mu_2\left[\mu_2^{-1}\mathcal{M}+t\varepsilon\kappa\frac{D^2u}{q+1}\right] \\ &= \mu_1\mu_2\left[\mathrm{Id}-\kappa\varepsilon^2 Du\otimes Du+t\varepsilon\kappa\frac{D^2u}{q+1}\right]. \end{aligned}$$

This finishes the proof of Proposition 4.1. □

4.2. **Proof of Theorem A.** Now we are ready to finish the proof of Theorem A. Let $u\in C^2(\mathcal{U})$ be a solution to the refractor problem (**RP**) then from Proposition 4.1 we obtain

$$\det Dz=\det\mu_1\det\mu_2\left[\frac{t\varepsilon\kappa}{q+1}\right]^n\det\left[\frac{q+1}{t\varepsilon\kappa}\left\{\mathrm{Id}-\kappa\varepsilon^2 Du\otimes Du\right\}+D^2u\right]. \tag{4.17}$$

By Lemma 3.2 and (1.10) we have

$$\det\mu_2=1+\frac{\kappa|Du|^2}{q(q+1)}=\frac{1-\kappa+q}{q(q+1)}.$$

Similarly, we get

$$\det\mu_1=\frac{\psi_{n+1}}{Y^{n+1}}\frac{1}{\nabla\psi\cdot Y}.$$

These in conjunction with (3.11) gives

$$\begin{aligned} \det\left[\frac{q+1}{t\varepsilon\kappa}\left\{\mathrm{Id}-\kappa\varepsilon^2 Du\otimes Du\right\}+D^2u\right] &= \left[\frac{q+1}{t\varepsilon\kappa}\right]^n\frac{\det Dz}{\det\mu_1\det\mu_2} \\ &= -\frac{f}{g}\frac{\psi_{n+1}}{|\nabla\psi|}\left[\frac{q+1}{t\varepsilon\kappa}\right]^n\frac{1}{\det\mu_1\det\mu_2} \\ &= -(\nabla\psi\cdot Y)\frac{Y^{n+1}}{|\nabla\psi|}\frac{q(q+1)}{1-\kappa+q}\left[\frac{q+1}{t\varepsilon\kappa}\right]^n\frac{f}{g}. \end{aligned}$$

Finally, recalling (3.8) and substituting the value of Y^{n+1} we see that

$$\begin{aligned} \det\left[\frac{q+1}{t\varepsilon\kappa}\left\{\mathrm{Id}-\kappa\varepsilon^2 Du\otimes Du\right\}+D^2u\right] &= -(\nabla\psi\cdot Y)\frac{Y^{n+1}}{|\nabla\psi|}\frac{q(q+1)}{1-\kappa+q}\left[\frac{q+1}{t\varepsilon\kappa}\right]^n\frac{f}{g} \\ &= -\varepsilon q\left[\frac{q+1}{t\varepsilon\kappa}\right]^n\frac{\nabla\psi\cdot Y}{|\nabla\psi|}\frac{f}{g} \end{aligned} \tag{4.18}$$

and the proof of Theorem A is now complete. □

5. Existence of smooth solutions

In this section we will have a provisional discussion on the existence of smooth solutions to (1.9). Our main objective is to apply the available regularity theory for the Monge-Ampère type equations, stemming from seminal paper [14], in order to establish the regularity of weak solutions of the refractor problem.

We first rewrite the equation (4.18) in a more concise form. Let us introduce the following matrix

$$G_{ij} = \frac{1}{t}(q+1)[\delta_{ij} - \kappa\varepsilon^2 u_i u_j]. \tag{5.1}$$

Here $q = \sqrt{1-\kappa(1+|Du|^2)}$, see (1.10) and t is the stretch function determined from implicit equation $\psi(x + e_{n+1}u + tY) = 0$ as in Theorem A. Then the equation (4.18) transforms into

$$\det\left[-\frac{G}{\varepsilon\kappa} - D^2u\right] = |h(x,u,Du)|, \quad \text{if } \kappa > 0, \varepsilon > 1 \quad u \in C^2(\mathcal{U}) \text{ and } -\frac{G}{\varepsilon\kappa} - D^2u \geq 0, \tag{5.2}$$

$$\det\left[D^2u + \frac{G}{\varepsilon\kappa}\right] = |h(x,u,Du)|, \quad \text{if } \kappa < 0, \varepsilon < 1 \quad u \in C^2(\mathcal{U}) \text{ and } \frac{G}{\varepsilon\kappa} + D^2u \geq 0 \tag{5.3}$$

with

$$h(x,u,Du) = -\varepsilon q\left[\frac{q+1}{t\varepsilon\kappa}\right]^n \frac{\nabla\psi\cdot Y}{|\nabla\psi|}\frac{f}{g}. \tag{5.4}$$

The existence of C^2 smooth solutions of (5.2) or (5.3) depend on the properties of the matrix G. Namely, it is shown in [14] that if we regard G as a function of variable $p = Du$ then the condition

$$D^2_{p_k p_l} G_{ij}\xi_i\xi_j\eta_k\eta_l \begin{array}{ll} \leq -c_0|\xi|^2|\eta|^2 & \text{if } \kappa > 0 \\ \geq c_0|\xi|^2|\eta|^2 & \text{if } \kappa < 0 \end{array} \qquad \forall \xi,\eta \in \mathbb{R}^n, \xi \perp \eta, \tag{5.5}$$

with c_0 being a positive constant, is sufficient to obtain a priori $C^{1,1}$ bounds for the smooth solutions.

It is noteworthy to point out that the condition (5.5) and the C^2 estimates were derived in [14] for the Monge-Ampère type equations with variational structure emerging in optimal transport theory. The method used there is based on comparison the weak solution with the smooth one in a small ball. To employ this method successfully in the outset of refractor problem we need to establish a comparison principle, suitable mollification of the weak solution and a priori estimated for the smooth solutions of Dirichlet's problem in small balls.

The method outlined above gives the C^2 estimates for non-variational case as well, see [9, 10]. Therefore the local regularity result for the solutions to (5.2)-(5.3) with smooth w will follow once the matrix G verifies the condition (5.5). That done, the regularity of weak solutions reduces to the verification of the inequality (5.5) with some positive constant c_0.

The conditions imposed on the matrix in (5.2)-(5.3) involving the Hessian implies that the Monge-Ampère equation is degenerate elliptic. The weak formulation of degenerate ellipticity will be discussed in Section 10. Postponing the precise definition of weak solutions until then we would like to point out how the ellipticity of equation follows if we consider those C^2 solutions of (5.2) (reap. (5.3)) for which at every point $x \in \mathcal{U}$ there is a hyperboloid (reap. ellipsoid) of revolution $H(\cdot, a, Z)$ touching u from above (reps. below) at x. Indeed, for $H(x) = \ell_0 - \frac{a}{b}\sqrt{b^2 + |x - x_0|^2}$ the matrix $\mathcal{W}_H = -\frac{G}{\varepsilon\kappa} + D^2H$ is identically zero. To see this we consider the case of planar receiver Σ given as $X^{n+1} = m$ with $m > 0$. Without loss of generality we take $x_0 = 0$. Then $H(0) = \ell_0 - a$. On the other hand it follows from the definition of eccentricity $\varepsilon = \frac{\sqrt{a^2+b^2}}{a}$ that $\ell_0 = m - a\varepsilon$, see Section 7. Next, a simple geometric reasoning yields the following explicit formula for the stretch function

$$t = \frac{m - H}{Y^{n+1}} = \frac{c + \frac{a}{b}\sqrt{b^2 + |x|^2}}{\varepsilon(1 - \frac{\kappa}{h+1})}. \tag{5.6}$$

We have $DH=-\frac{a}{b}\frac{x}{\sqrt{b^2+|x|^2}}$. Consequently

$$D^2H=-\frac{a}{b\sqrt{b^2+|x|^2}}\left(\mathrm{Id}-\frac{x\otimes x}{b^2+|x|^2}\right). \tag{5.7}$$

Moreover, recalling (3.3) we obtain $\kappa=1-\frac{1}{\varepsilon^2}=\frac{b^2}{c^2}$ where $c=\sqrt{a^2+b^2}$. This gives

$$q(x)=\frac{1}{\varepsilon}\frac{b}{\sqrt{b^2+|x|^2}}, \tag{5.8}$$

in lieu of (1.10).

Thus combining these formulae for t and q we get from (5.1), (5.6) and (5.7)

$$\begin{aligned}\mathcal{W}_H &= -\frac{G}{\varepsilon\kappa}-D^2H\\ &= -\frac{q+1}{\kappa\varepsilon t}\left\{[\delta_{ij}-\kappa\varepsilon^2H_iH_j]+\frac{\kappa\varepsilon tD^2_{ij}H}{q+1}\right\}\\ &= -\frac{q+1}{\kappa\varepsilon t}\left\{\mathrm{Id}-\underbrace{\left[\kappa\varepsilon^2\frac{a^2}{b^2}\right]\frac{x\otimes x}{b^2+|x|^2}}_{DH\otimes DH}+\frac{\kappa\varepsilon t}{q+1}\underbrace{\left[-\frac{a}{b\sqrt{b^2+|x|^2}}\left(\mathrm{Id}-\frac{x\otimes x}{b^2+|x|^2}\right)\right]}_{D^2H}\right\}.\end{aligned}$$

From the definition of κ (1.7) it follows that $\kappa\varepsilon^2\frac{a^2}{b^2}=\frac{b^2}{c^2}\varepsilon^2\frac{a^2}{b^2}=1$ implying

$$\mathcal{W}_H=-\frac{q+1}{\kappa\varepsilon t}\left(\mathrm{Id}-\frac{x\otimes x}{b^2+|x|^2}\right)\left\{1-\frac{\kappa\varepsilon t}{q+1}\left[\frac{a}{b\sqrt{b^2+|x|^2}}\right]\right\}.$$

Therefore, recalling (5.6) and (5.8) we easily compute

$$\begin{aligned}t &= \frac{c+\frac{a}{b}\sqrt{b^2+|x|^2}}{\varepsilon(1-\frac{\kappa}{h+1})}=(q+1)\frac{c+\frac{a}{b}\sqrt{b^2+|x|^2}}{\varepsilon(q+1-\kappa)}\\ &= (q+1)\frac{c+\frac{a}{b}\sqrt{b^2+|x|^2}}{\varepsilon(q+\frac{1}{\varepsilon^2})}\\ &= \varepsilon(q+1)\frac{c+\frac{a}{b}\sqrt{b^2+|x|^2}}{\varepsilon^2q+1}\\ &= \varepsilon(q+1)\sqrt{b^2+|x|^2}\frac{c+\frac{a}{b}\sqrt{b^2+|x|^2}}{\varepsilon b+\sqrt{b^2+|x|^2}}.\end{aligned} \tag{5.9}$$

Returning to $\mathcal{W}_H$ and utilizing (5.9) we obtain

$$\begin{aligned}1-\frac{\kappa\varepsilon t}{q+1}\left[\frac{a}{b\sqrt{b^2+|x|^2}}\right] &= 1-\kappa\varepsilon^2\frac{c+\frac{a}{b}\sqrt{b^2+|x|^2}}{\sqrt{b^2+|x|^2}+b\varepsilon}\\ &= 1-\varepsilon\kappa\frac{c}{a}\frac{a^2}{b^2}\\ &= 1-\varepsilon^2\kappa\frac{a^2}{b^2}\\ &= 0.\end{aligned}$$

A similar computation for the matrix $\mathcal{W}_E=\dfrac{G}{t\varepsilon|\kappa|}+D^2E$ can be carried out for the ellipsoids of revolution E (i.e. for $\varepsilon<1,\kappa<0$).

Since $-D^2u \geq -D^2H_{x_0}$ at x_0 and $\mathcal{W}_H = -\frac{G}{\varepsilon\kappa} - D^2H_{x_0} \equiv 0$ it follows that the equation $\det\left[-\frac{G}{\varepsilon\kappa} - D^2u\right] = h$ is degenerate elliptic.

Notice that for $\varepsilon < 1$ the weak solution has a supporting ellipsoid of revolution E_{x_0} at each point $x_0 \in \overline{\mathcal{U}}$ touching Γ_u from below. In particular we see that if $u \in C^2$ then $Du = DE_{x_0}, -D^2u \leq -D^2E_{x_0}$ at x_0. Thus $\dfrac{G}{\varepsilon\kappa} + D^2u \geq 0$ and we infer that (5.2) is degenerate elliptic. Analogously, using the hyperboloids as supporting functions, one can check that (5.3) is also degenerate elliptic.

6. Proof of Theorem B: Verifying the A3 condition

In this section we explicitly compute $D_{p_kp_l}G^{ij}\xi^i\xi^j\eta^k\eta^l$ explicitly and relate it with the second fundamental form of the receiver $\Sigma = \{Z \in \mathbb{R}^{n+1} \ : \ \psi(Z) = 0$ where $\psi : \mathbb{R}^{n+1} \to \mathbb{R}$ is a smooth function such that (1.11) holds.

6.1. **Computing the derivatives of stretch function t.** Recall that by (3.10) $Z(x) = x + e_{n+1}u(x) + tY$. Differentiating $\psi(Z(x)) = 0$ with respect to p_k we get

$$\frac{t_{p_k}}{t} = -\frac{\sum_m \psi_m Y^m_{p_k}}{\sum_m \psi_m Y^m}, \quad k = 1, \dots n. \tag{6.1}$$

After differentiating again by p_l we get

$$\begin{aligned}
\frac{t_{p_kp_l}}{t} - \frac{t_{p_k}t_{p_l}}{t^2} &= -\Bigg[\frac{\sum_{ms}\psi_{ms}(Y^s_{p_l}t + Y^st_{p_l})Y^m_{p_k} + \sum \psi_m Y^m_{p_kp_l}}{(\nabla\psi\cdot Y)} \\
&\quad -\frac{\sum_m \psi_m Y^m_{p_k}}{(\nabla\psi\cdot Y)^2}\left(\sum_{m,s}\psi_{ms}(Y^s_{p_l}t + Y^st_{p_l})Y^m + \sum_m \psi_m Y^m_{p_l}\right)\Bigg] \\
&= -\frac{1}{(\nabla\psi\cdot Y)}\Bigg[\left(\nabla^2\psi Y_{p_k}Y_{p_l} - \frac{\nabla\psi Y_{p_k}}{(\nabla\psi\cdot Y)}\nabla^2\psi Y Y_{p_l}\right)t \\
&\quad + \left(\nabla^2\psi Y_{p_k}Y - \frac{\nabla\psi Y_{p_k}}{(\nabla\psi\cdot Y)}\nabla^2\psi YY\right)t_{p_l} \\
&\quad + \nabla\psi Y_{p_kp_l} - \frac{\nabla\psi Y_{p_k}}{(\nabla\psi\cdot Y)}\nabla\psi Y_{p_l}\Bigg] = \\
&= -\frac{1}{(\nabla\psi\cdot Y)}\Bigg[\left(\nabla^2\psi Y_{p_k}Y_{p_l} - \frac{\nabla\psi Y_{p_k}}{(\nabla\psi\cdot Y)}\nabla^2\psi Y Y_{p_l}\right)t \\
&\quad + \left(\nabla^2\psi Y_{p_k}Y - \frac{\nabla\psi Y_{p_k}}{(\nabla\psi\cdot Y)}\nabla^2\psi YY\right)t_{p_l} + \nabla\psi Y_{p_kp_l}\Bigg] \\
&\quad + \frac{t_{p_k}t_{p_l}}{t^2}.
\end{aligned} \tag{6.2}$$

Rearranging the terms we infer

$$\begin{aligned}
D^2_{p_kp_l}\left(\frac{1}{t}\right) = -\frac{t_{p_kp_l}}{t^2} + \frac{2t_{p_k}t_{p_l}}{t^3} &= \frac{1}{t(\nabla\psi\cdot Y)}\Bigg[\left(\nabla^2\psi Y_{p_k}Y_{p_l} - \frac{\nabla\psi Y_{p_k}}{(\nabla\psi\cdot Y)}\nabla^2\psi YY_{p_l}\right)t \\
&\quad + \left(\nabla^2\psi Y_{p_k}Y - \frac{\nabla\psi Y_{p_k}}{\nabla\psi\cdot Y}\nabla^2\psi YY\right)t_{p_l} + \nabla\psi Y_{p_kp_l}\Bigg] \\
&= \frac{1}{t(\nabla\psi\cdot Y)}\left[\frac{1}{t}\nabla^2\psi Z_{p_k}Z_{p_l} + \nabla\psi Y_{p_kp_l}\right]
\end{aligned} \tag{6.3}$$

where the last line follows from (6.1). Thus the second derivatives of $\frac{1}{t}$ can be computed from (6.3), while for the first order derivatives we have the formula (6.1).

Next, we want to compute the derivatives of $M_{ij} = (q+1)[\delta_{ij} - \kappa\varepsilon^2 p_ip_j]$ with respect to p. We have

$$D_{p_k}M_{ij} = q_{p_k}(\delta_{ij} - \kappa\varepsilon^2 p_ip_j) - (q+1)\kappa\varepsilon^2[\delta_{kj}p_i + \delta_{ki}p_j].$$

The condition $\xi\perp\eta$ implies that the contribution of the terms involving δ_{kj} and δ_{ki} is zero. Thus we infer

$$D^2_{p_kp_l}M_{ij}\xi^i\xi^j\eta^k\eta^l=q_{p_kp_l}\eta^k\eta^l\left[|\xi|^2-\kappa\varepsilon^2(p\cdot\xi)^2\right].$$

Recall that by definition $G=\frac{M}{t}$ hence from the product rule we have

$$\begin{aligned}
D^2_{p_kp_l}G^{ij}\xi^i\xi^j\eta^k\eta^l &= D^2_{p_kp_l}\left(\frac{1}{t}\right)\eta^k\eta^l(q+1)\left[|\xi|^2-\kappa\varepsilon^2(p\cdot\xi)^2\right]\\
&\quad+2\left(D_p\frac{1}{t}\cdot\eta\right)(D_pq\cdot\eta)\left[|\xi|^2-\kappa\varepsilon^2(p\cdot\xi)^2\right]\\
&\quad+\frac{1}{t}(D^2_{p_kp_l}q)\eta^k\eta^l\left[|\xi|^2-\kappa\varepsilon^2(p\cdot\xi)^2\right]\\
&= S_{kl}\eta^k\eta^l\left[|\xi|^2-\kappa\varepsilon^2(p\cdot\xi)^2\right],
\end{aligned}\tag{6.4}$$

where

$$S_{kl}=(q+1)D^2_{p_kp_l}\left(\frac{1}{t}\right)+D_{p_k}\left(\frac{1}{t}\right)D_{p_l}q+D_{p_l}\left(\frac{1}{t}\right)D_{p_k}q+\frac{D^2_{p_kp_l}q}{t}.\tag{6.5}$$

It follows from (3.8) that

$$Y_{p_kp_l}(q+1)+Y_{p_k}q_{p_l}+Y_{p_l}q_{p_k}+Yq_{p_kp_l}=e_{n+1}\varepsilon q_{p_kp_l}.\tag{6.6}$$

which after taking the inner product with $\nabla\psi$ and dividing the by $\nabla\psi\cdot Y$ yields

$$\begin{aligned}
\frac{(q+1)\psi_{n+1}q_{p_kp_l}}{\nabla\psi\cdot Y} &= \frac{(q+1)\nabla\psi\cdot Y_{p_kp_l}}{\nabla\psi\cdot Y}+\frac{q_{p_k}\nabla\psi\cdot Y_{p_l}}{\nabla\psi\cdot Y}+\frac{q_{p_l}\nabla\psi\cdot Y_{p_k}}{\nabla\psi\cdot Y}+q_{p_kp_l}=\\
&= \frac{(q+1)\nabla\psi\cdot Y_{p_kp_l}}{\nabla\psi\cdot Y}+\frac{q_{p_k}\nabla\psi\cdot Y_{p_l}}{\nabla\psi\cdot Y}+\frac{q_{p_l}\nabla\psi\cdot Y_{p_k}}{\nabla\psi\cdot Y}+q_{p_kp_l}\\
&= \frac{(q+1)\nabla\psi\cdot Y_{p_kp_l}}{\nabla\psi\cdot Y}+tq_{p_k}D_{p_l}\left(\frac{1}{t}\right)+tq_{p_l}D_{p_k}\left(\frac{1}{t}\right)+q_{p_kp_l}\\
&= \frac{(q+1)\nabla\psi\cdot Y_{p_kp_l}}{\nabla\psi\cdot Y}+t\left(S_{kl}-(q+1)D^2_{p_kp_l}\left(\frac{1}{t}\right)\right).
\end{aligned}\tag{6.7}$$

Consequently, with the aid of (6.3) we find that

$$\begin{aligned}
S_{kl} &= \frac{q+1}{t(\nabla\psi\cdot Y)}\left(\psi_{n+1}q_{p_kp_l}-\nabla\psi\cdot Y_{p_kp_l}\right)+(q+1)D^2_{p_kp_l}\left(\frac{1}{t}\right)\\
&= \frac{q+1}{t(\nabla\psi\cdot Y)}\left(\psi_{n+1}q_{p_kp_l}-\nabla\psi\cdot Y_{p_kp_l}\right)+\frac{q+1}{t(\nabla\psi\cdot Y)}\left[\frac{1}{t}\nabla^2\psi Z_{p_k}Z_{p_l}+\nabla\psi Y_{p_kp_l}\right]\\
&= \frac{q+1}{t(\nabla\psi\cdot Y)}\left[\frac{1}{t}\nabla^2\psi Z_{p_k}Z_{p_l}+\psi_{n+1}q_{p_kp_l}\right].
\end{aligned}$$

It remains to recall that by (1.10)

$$q_{p_kp_l}=-\frac{\kappa}{q}\left[\delta_{kl}+\kappa\frac{p_kp_l}{q^2}\right]\tag{6.8}$$

and we conclude

$$D^2_{p_kp_l}G^{ij}\xi^i\xi^j\eta^k\eta^l=\frac{q+1}{t(\nabla\psi\cdot Y)}\left[\frac{1}{t}\nabla^2\psi Z_{p_k}Z_{p_l}+\psi_{n+1}q_{p_kp_l}\right]\eta^k\eta^l\left[|\xi|^2-\kappa\varepsilon^2(p\cdot\xi)^2\right].\tag{6.9}$$

It is worth noting that $|\xi|^2-\kappa\varepsilon^2(p\cdot\xi)^2$ is always positive. This is obvious if $\kappa<0$. As for $\kappa>0$ then we note that $|\xi|^2-\kappa\varepsilon^2(p\cdot\xi)^2=|\xi|^2\left(1-\kappa\varepsilon^2\left(p\cdot\frac{\xi}{|\xi|}\right)^2\right)>0$ in view of the estimate $|p|\le\frac{1}{\sqrt{\varepsilon^2-1}}$, see (1.16). Furthermore, from (1.11) it follows that $D^2_{p_kp_l}G^{ij}$ and $\widehat{S}_{kl}$ defined by

$$\widehat{S}_{kl}=\frac{1}{t}\nabla^2\psi Z_{p_k}Z_{p_l}+\psi_{n+1}q_{p_kp_l}\tag{6.10}$$

have the same signs. Thus it is enough to explore the form $\widehat{S}_{kl}$ instead.

6.2. **Refining condition (5.5).** Let Z_0 be a fixed point on Σ. Introduce a new coordinate system $\hat{x}_1, \dots, \hat{x}_n, \hat{x}_{n+1}$ near Z_0, with $\hat{x}_{n+1}$ having direction Y. Since (1.11) and (1.12) implies $\nabla\psi \neq 0$, without loss of generality we assume that near Z_0, in $\hat{x}_1, \dots, \hat{x}_n, \hat{x}_{n+1}$ coordinate system Σ has a representation $\hat{x}_{n+1} = \varphi(\hat{x}_1, \dots, \hat{x}_n)$. Recall that the second fundamental form of Σ is

$$\mathrm{II} = \frac{\partial^2_{\hat{x}_i, \hat{x}_j}\varphi}{\sqrt{1+|D\varphi|^2}}, \qquad i,j = 1, \dots, n \tag{6.11}$$

if we choose the normal of Σ at Z_0 to be $\frac{(-D_{\hat{x}_1}\varphi, \dots, -D_{\hat{x}_n}\varphi, 1)}{\sqrt{1+|D\varphi|^2}}$, $D\varphi = (D_{\hat{x}_1}\varphi, \dots, D_{\hat{x}_n}\varphi, 0)$.

Denote $\widetilde{\psi}(Z) = Z^{n+1} - \varphi(z)$ and assume that near Z_0, Σ is given by the equation $\widetilde{\psi} = 0$. It follows that

$$\nabla^2\widetilde{\psi} = -\begin{vmatrix} \varphi_{11} & \cdots & \varphi_{1n} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \varphi_{n1} & \cdots & \varphi_{nn} & 0 \\ 0 & \cdots & 0 & 0 \end{vmatrix}. \tag{6.12}$$

Therefore for $Z = x + ue_{n+1} + tY$ we have $\nabla^2\widetilde{\psi}Y = 0$ and hence

$$\begin{aligned} \nabla^2\widetilde{\psi}Z_{p_k}Z_{p_l} &= \nabla^2\widetilde{\psi}(tY_{p_k} + t_{p_k}Y)(tY_{p_l} + t_{p_l}Y) \\ &= t^2\nabla^2\widetilde{\psi}Y_{p_k}Y_{p_l} \\ &= -t^2\nabla^2\varphi Y_{p_k}Y_{p_l}. \end{aligned} \tag{6.13}$$

By (4.12) $Y(q+1) = (\varepsilon\kappa p, \varepsilon q + \varepsilon - \kappa)$ where $y(q+1) = \varepsilon\kappa p$. Differentiating this equality with respect to p_k we infer

$$Y_{p_k}(q+1) + Yq_{p_k} = \varepsilon(\kappa\hat{e}_k + q_{p_k}\hat{e}_{n+1}) \tag{6.14}$$

hence

$$Y_{p_k} = \frac{1}{q+1}\left[-Yq_{p_k} + \varepsilon(\kappa\hat{e}_k + q_{p_k}\hat{e}_{n+1})\right]. \tag{6.15}$$

On the other hand (6.12) and $\hat{e}_{n+1} = Y$ yield

$$\begin{aligned} \nabla^2\widetilde{\psi}Y_{p_k} &= \frac{1}{q+1}\nabla^2\widetilde{\psi}\left[-Yq_{p_k} + \varepsilon(\kappa\hat{e}_k + q_{p_k}\hat{e}_{n+1})\right] \\ &= \frac{\varepsilon}{q+1}\nabla^2\widetilde{\psi}(\kappa\hat{e}_k + q_{p_k}\hat{e}_{n+1}) \\ &= \frac{\varepsilon\kappa}{q+1}\nabla^2\widetilde{\psi}\hat{e}_k. \end{aligned} \tag{6.16}$$

Since $\nabla^2\widetilde{\psi}$ is symmetric we infer

$$\nabla^2\widetilde{\psi}Y_{p_k}Y_{p_l} = \frac{\varepsilon^2\kappa^2}{(q+1)^2}\nabla^2\widetilde{\psi}\hat{e}_k\hat{e}_l. \tag{6.17}$$

Plugging (6.17) into (6.13) we finally obtain

$$\nabla^2\widetilde{\psi}Z_{p_k}Z_{p_l} = -t^2\frac{\varepsilon^2\kappa^2}{(q+1)^2}\nabla^2\varphi\hat{e}_k\hat{e}_l. = -t^2\frac{\varepsilon^2\kappa^2}{(q+1)^2}\sqrt{1+|D\varphi|^2}\mathrm{II} \tag{6.18}$$

where II is the second fundamental form of Σ at Z_0, see (6.11). This in conjunction with (6.8) yields

$$\widehat{S}_{kl} = -\left[\frac{1}{t}\left(\frac{t\varepsilon\kappa}{q+1}\right)^2\sqrt{1+|D\varphi|^2}\mathrm{II}+\frac{\kappa}{q}(\mathrm{Id}+\kappa\frac{p\otimes p}{q^2})\right]. \tag{6.19}$$

6.3. **Planar receivers.** Let us consider the case of horizontal receiver $Z^{n+1} = m > 0$ for some positive number m. Then $\mathrm{II} = 0$ implying that $\widehat{S}_{kl} = -\frac{\kappa}{q}(\mathrm{Id}+\kappa\frac{p\otimes p}{q^2})$. If $\kappa > 0$ then clearly $\widehat{S}_{kl} < -\frac{\kappa}{q}\delta_{kl} \le -c_0\delta_{kl}$, where $c_0 > 0$ depends only on $\sup|p|$ and ε. As for $\kappa < 0$ we compute

$$S_{kl}\xi^i\xi^j = \frac{|\kappa||\xi|^2}{q}(1-\frac{|\kappa|(p\cdot\xi)^2}{q^2|\xi|^2}) \ge \frac{|\kappa||\xi|^2}{q}(1-\frac{|\kappa|p|^2}{q^2}) = \frac{|\kappa|(1+|\kappa|)}{q^3}|\xi|^2 \ge c_0|\xi|^2$$

where $c_0 > 0$ depends only on $\sup|p|$ and ε. Consequently (5.5) is true for horizontal receivers $Z^{n+1} = m > 0$.

Remark 6.1. *The computation above shows that (5.5) is true if $\kappa > 0, \mathrm{II} \ge 0$ or if $\kappa < 0, \mathrm{II} \le 0$. We can extend Σ to entire space such that the resultd surface is still concave if say $\kappa > 0$, hence without loss of generality we can assume that Σ is entire concave surface and so is $\Sigma + Me_{n+1}$, for $M \gg 1$. We will take advantage of this in Lemmas 7.3 and 9.2*

7. Admissible functions

The refractive properties of ellipses and hyperbolas have been known since ancient times [15]. Furthermore, hyperboloids and ellipsoids of revolution share the same properties. This section is devoted to the class of functions obtained as envelopes of halves of ellipsoids and hyperboloids of revolution.

7.1. **Ellipsoids.** Throughout this paper by ellipsoid we mean the lower half of an ellipsoid of revolution with focal axis parallel to e_{n+1}. Such surface can be regarded as the graph of

$$E(x,a,Z) = Z^{n+1} - a\varepsilon - a\sqrt{1-\frac{(x-z)^2}{a^2(1-\varepsilon^2)}} \tag{7.1}$$

where a is the larger semiaxis, ε- the eccentricity, and Z the higher focus, see Figure 2. Moreover we have that

$$DE = \frac{1}{a(1-\varepsilon^2)}\frac{x-z}{\sqrt{1-\frac{(x-z)^2}{a^2(1-\varepsilon^2)}}}. \tag{7.2}$$

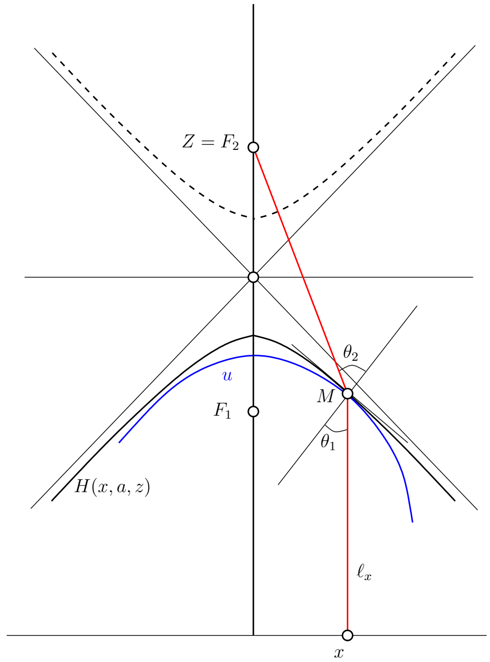
Notice that at the points x where $|x-z| = a\sqrt{1-\varepsilon^2}$ the gradient $|DE|$ is unbounded.

7.2. **Hyperboloids.** It is convenient to introduce the lower sheet of hyperboloids of revolution

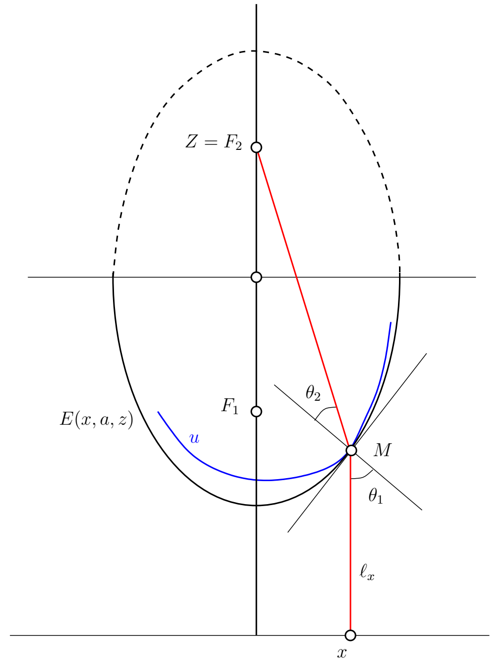
$$H(x,a,Z) = Z^{n+1} - a\varepsilon - a\sqrt{1+\frac{(x-z)^2}{a^2(\varepsilon^2-1)}} \tag{7.3}$$

where a is the larger semiaxis, ε the eccentricity, and Z the upper focus, see Figure 2. Differentiating H we obtain

$$DH = -\frac{1}{(\varepsilon^2-1)}\frac{x-z}{\sqrt{a^2+\frac{(x-z)^2}{\varepsilon^2-1}}}. \tag{7.4}$$



(A) Upper admissible



(B) Lower admissible

FIGURE 2. Supporting surfaces

7.3. **Supporting hyperboloids.**

Definition 7.1. *A function $u : \mathcal{U} \to \mathbb{R}$ is said to be upper (resp. lower) admissible if for any $x_0 \in \mathcal{U}$ there is $Z \in \mathcal{V}$ and $a > 0$ such that $H(x_0, a, Z) = u(x_0)$ (resp. $E(x_0, a, Z) = u(x_0)$) and $H(x, a, Z) \geq u(x), x \in \mathcal{U}$ (resp. $E(x, a, Z) = u(x)$). H (resp. E) is called a supporting function of u at x_0. The class of all upper admissible functions is denoted by $\overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \mathcal{V})$ (resp. $\underline{\mathbb{W}}_{\mathrm{E}}(\mathcal{U}, \mathcal{V})$).*

In what follows we focus on upper admissible functions, the lower admissible functions can be studied in similar fashion. If the generalisation is not straightforward then we will outline the proof.

Formula (7.4) yields uniform Lipschitz estimates for $\overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \mathcal{V})$.

Lemma 7.1. *Let $\mathbb{H}^+(\mathcal{U}, \mathcal{V})$ be the set of all hyperboloids $H(x, a, Z) \geq 0, Z \in \mathcal{V}$ for any $x \in \mathcal{U}$. Then*

$$\sup_{\mathbb{H}^+(\mathcal{U}, \mathcal{V})} \|DH\|_\infty < \frac{1}{\sqrt{\varepsilon^2 - 1}}.$$

In particular

$$\sup_{\overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \mathcal{V})} \|Du\|_\infty < \frac{1}{\sqrt{\varepsilon^2 - 1}}.$$

Proof. From (7.4) we have

$$|DH| = \frac{1}{(\varepsilon^2 - 1)} \frac{|x - z|}{\sqrt{a^2 + \frac{|x-z|^2}{\varepsilon^2 - 1}}} = \frac{1}{\sqrt{\varepsilon^2 - 1}} \frac{|x - z|}{\sqrt{a^2(\varepsilon^2 - 1) + |x - z|^2}}.$$

Since $|x - z|$ is uniformly bounded, the result follows. $\square$

Lemma 7.2. *Let $\{u_k\}$ be a sequence of upper admissible function such that $u_k \to u_0$ uniformly in $\mathcal{U}$. If $x_k \in \mathcal{U}$, $x_k \to x_0$ and H_k are supporting functions of u_k at x_k then u_0 has an upper supporting function H_0 at x_0 and $H_k \to H_0$ uniformly in $\mathcal{U}$.*

Proof. One way to check the claim is to use some well known fact from convex analysis. Consider the convex sets $G_k = \{X \in \mathbb{R}^{n+1} : x \in \mathcal{U}, 0 < u_k(x) < X^{n+1}\}$ and $\mathscr{H}_k = \{X \in \mathbb{R}^{n+1} : x \in \mathcal{U}, 0 < H_k(x) < X^{n+1}\}$ where $x = \widehat{X}$. Then $\mathscr{H}_k \subset G_k$ and $(x_k, u_k(x_k)) \in \partial\mathscr{H}_k \cap \partial G_k$. Thus, from uniform convergence $u_k \to u_0$ we infer that the limit set $G_0 = \{X \in \mathbb{R}^{n+1} : x \in \mathcal{U}, 0 < u_0(x) < X^{n+1}\}$ is a subset of $\mathscr{H}_0 = \{X \in \mathbb{R}^{n+1} : x \in \mathcal{U}, 0 < H_0(x) < X^{n+1}\}$, see [1] Chapter 5.2. Furthermore, from $x_k \to x_0 \in \overline{\mathcal{U}}$ it follows that there is $X_0 \in \partial G_0 \cap \partial\mathscr{H}_0$ such that $\widehat{X}_0 = x_0$. Therefore we conclude that H_0 is a supporting hyperboloid of u_0 at x_0. □

7.4. **Continuous expansion of hyperboloids.** If $u \in \overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \Sigma)$ then it turns out that u is also admissible with respect with $\widetilde{\Sigma}$, the receiver moved vertically upwards in e_{n+1} direction. In other words, the same admissible u will be $R-$convex with respect to a family of surfaces obtained from Σ by translation is e_{n+1} direction. We will need this observation in order to construct smooth solutions of our problem in small balls, see Section 13.

Lemma 7.3. *Let $\widetilde{\Sigma} = \Sigma + Me_{n+1}$ for some $M > 0$.*

(i) *For any fixed x_0 and $H_1(x) = H(x, a_1, Z_1) \in \mathbb{H}(\mathcal{U}, \Sigma)$ there is $H_2(x) = H(x, a_2, Z_2)$ with $Z_2 \in \widetilde{\Sigma}$ and touching H_1 from above at x_0.*

(ii) *In particular if $u \in \overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \Sigma)$ then also $u \in \overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \widetilde{\Sigma})$.*

Proof. (i) Let $\xi_1 = H_1(x_0)$ and $X_0 = (x_0, \xi_1)$. For $s > 1$ we consider $Z_2 = X_0 + s(Z_1 - X_0)$. By construction X_0, Z_1 and Z_2 lie on the same line. To determine a_2 we utilize two geometric properties of hyperbola, namely that the difference of distances of X_0 from Z_2 and the lower focus Z_2' is $2a_2$ and $|X_0Z_2'| = \varepsilon|X_0D|$ where $|X_0D|$ is the distance of X_0 from the lower directrix $X^{n+1} = Z^{n+1} - a_2\varepsilon - a_2/\varepsilon$. Therefore if P is on the graph of H_2 we get that $|PZ_2| = 2a_2 + |PZ_2'| = -a_2(\varepsilon^2 - 1) + s\varepsilon(Z_1^{n+1} - \xi_1)$. Taking $P = X_0$ in this equation $|PZ_2| = \varepsilon|Z_1 - X_0|$ one finds that

$$a_2 = \frac{1}{\varepsilon^2 - 1}[s\varepsilon(Z_1^{n+1} - \xi_1) - |s(Z_1 - X_0)|]. \tag{7.5}$$

As for (ii), we choose $s_0 > 1$ so that $X_0 + s(Z_1 - X_0) \in \widetilde{\Sigma}$. Consequently from (i) it follows that $Z_2 = X_0 + s_0(Z_1 - X_0)$ is the focus of supporting hyperboloid $H(\cdot, a_2, Z_2)$ at x_0 where a_2 is given by (7.5). Therefore $u \in \overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \widetilde{\Sigma})$. □

8. B-type weak solutions: Proof of Theorem C1

In this section we introduce our first notion of weak solution for the refractor problem (**RP**). For any upper admissible function $u \in \overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \mathcal{V})$ we define the mapping $\mathscr{S}_u : \mathcal{V} \to \mathcal{U}$ as follows

$$\mathscr{S}_u(Z) = \{x \in \mathcal{U} : \exists \text{ a supporting hyperboloid of } u \text{ at } x \text{ with focus at } Z \in \mathcal{V}\}.$$

For any Borel set $\omega \subset \mathcal{V}$ we put

$$\mathscr{S}_u(\omega) = \bigcup_{Z \in \omega} \mathscr{S}_u(Z). \tag{8.1}$$

We will write $\mathscr{S}(E)$ instead of $\mathscr{S}_u(E)$ if there is no confusion.

Proposition 8.1. *For $u \in \overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \mathcal{V})$ the corresponding mapping $\mathscr{S}$ enjoys the following properties:*

a) *$\mathscr{S}:\mathcal{V}\longrightarrow\Pi$ maps the closed sets to closed sets.*

b) *The mapping $\mathscr{S}$ is one-to-one modulo a set of vanishing measure, i.e.*

$$\left|\left\{x\in\Pi : x\in\mathscr{S}(Z_1)\cap\mathscr{S}(Z_2) \text{ for } Z_1\neq Z_2,\quad Z_i\in\mathcal{V}, i=1,2\right\}\right|=0.$$

c) *The family $\mathscr{F}=\{E\subset\mathcal{V}$ such that $\mathscr{S}(E)$ is measurable$\}$ is $\sigma-$algebra.*

Proof. The first claim a) follows directly from Lemma 7.2.

In order to prove b) we set $A=\left\{x\in\Pi : x\in\mathscr{S}(Z_1)\cap\mathscr{S}(Z_2) \text{ for } Z_1\neq Z_2,\quad Z_i\in\mathcal{V}, i=1,2\right\}$. If $x\in A$ then u cannot be differentiable at x. By Aleksandrov's theorem the concave function u is twice differentiable a.e. Hence $|A|=0$.

As for c) we must check that the following three conditions hold, see e.g. [2]

1) $\mathcal{V}\in\mathscr{F}$,
2) if $A\in\mathscr{F}$ then $\mathcal{V}\setminus A\in\mathscr{F}$,
3) if $A_i\in\mathscr{F}$ then $\bigcup_{i=1}^\infty A_i\in\mathscr{F}$.

We first prove 1). If $A_i\in\mathcal{V}$ is any sequence of subsets of $\mathcal{V}$ then clearly $\mathscr{S}(\cup_{i=1}^\infty A_i)=\cup_{i=1}^\infty\mathscr{S}(A_i)$. Writing $\mathcal{V}=\cup_{i=1}^\infty E_i$, where $E_i\subset\mathcal{V}$ are closed subsets we conclude that $\mathscr{S}(\mathcal{V})=\mathscr{S}(\cup_{i=1}^\infty E_i)=\cup_{i=1}^\infty\mathscr{S}(E_i)$. From a) it follows that $\mathscr{S}(E_i)$ is closed for any i, and hence measurable, implying that $\mathscr{S}(\mathcal{V})$ is measurable.

2) Let $A\in\mathscr{F}$. We use the following elementary identity

$$\mathscr{S}(\mathcal{V}\setminus A)=[\mathscr{S}(\mathcal{V})\setminus\mathscr{S}(A)]\bigcup[\mathscr{S}(\mathcal{V}\setminus A)\cap\mathscr{S}(A)]. \tag{8.2}$$

From b) it follows that $|\mathscr{S}(\mathcal{V}\setminus A)\cap\mathscr{S}(A)|=0$. Therefore $|\mathscr{S}(\mathcal{V}\setminus A)|=|\mathscr{S}(\mathcal{V})\setminus\mathscr{S}(A)|$ and 2) is proven.

It remains to check 3). Without loss of generality we assume that A_i's are disjoint, see [2]. Thus, letting $A_i\in\mathscr{F}, A_i\cap A_j=\emptyset, i\neq j$ we get

$$\begin{aligned}\sum_{i=1}^\infty|\mathscr{S}(A_i)| &\geq |\mathscr{S}(\cup_{i=1}^\infty A_i)|\geq\\ &\geq\sum_{i=1}^\infty|\mathscr{S}(A_i)|-\sum_{ij=1}^\infty|\mathscr{S}(A_i)\cap\mathscr{S}(A_j)|\geq\\ &\geq\sum_{i=1}^\infty|\mathscr{S}(A_i)|.\end{aligned}$$

□

For a given function $u\in\mathbb{W}_H(\mathcal{U},\mathcal{V})$ we consider the set function

$$\beta_u(\omega)=\int_{\mathscr{S}(\omega)}f \tag{8.3}$$

where $\omega\subset\mathcal{V}$ is a Borel subset. Since $\mathscr{F}$ contains the closed sets (see part a) above) we infer that $\beta_{u,f}$ is a Borel measure. Moreover, from the proof of Proposition 8.1 b) it follows that $\beta_{u,f}$ is countably additive.

Definition 8.1. *A function u (or its graph Γ_u) is said to be a B-type weak solution to (**RP**) if $u\in\overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U},\mathcal{V})$ and the following two identities holds*

$$\begin{cases}\beta_{u,f}(\omega)=\displaystyle\int_\omega g\,d\mathcal{H}^n, \text{ for any Borel set } \omega\subset\mathcal{V} \qquad \text{and}\\ \qquad\qquad\mathscr{S}_u(\mathcal{V})=\mathcal{U}.\end{cases} \tag{8.4}$$

8.1. **Existence of weak solutions of B-type.** The measure β, defined in (8.3) is weakly continuous. We have

Lemma 8.1. *Let u_k be a sequence of B-type weak solutions in the sense of Definition 8.3 and β_k is the associated measure, defined by (8.3). If $u_k \to u$ uniformly on compact subsets of $\mathcal{U}$ then u is $R-$concave and β_k weakly converges to $\beta_{u,f}$.*

Proof. That u is admissible follows from Lemma 7.2. Recall that the weak convergence is equivalent to the following two inequalities (see [2] Theorem 4.5.1)

1) $\limsup_{k\to\infty} \beta_k(E) \le \beta(E)$ for any compact $E \subset \mathcal{V}$,
2) $\liminf_{k\to\infty} \beta_k(J) \ge \beta(J)$ for any open $J \subset \mathcal{V}$.

Take a closed set E and let E^*_δ be an $\delta-$neighbourhood of the closed set $E^* = \mathscr{S}(E)$, see Lemma 8.1 a). We claim that for any $\delta > 0$ there is $i_0 \in \mathbb{N}$ such that $\mathscr{S}_i(E) \subset E^*_\delta$ whenever $i > i_0$, where $\mathscr{S}_i$ is the mapping corresponding to u_i. If this fails then there is $\delta > 0$ and a sequence of points $x_i \in \mathscr{S}_i(E)$ such that $x_i \in \complement E^*_\delta$. By definition there is $Z_i \in E$ such that $x_i \in \mathscr{S}_i(Z_i)$. Suppose that $x_i \to x_0$, for some x_0, and $Z_i \to Z_0 \in E$ at least for a subsequence. Thus, $x_0 \in \complement E^*_\delta, x_0 \in \mathscr{S}(Z_0)$ and $Z_0 \in E$ which is a contradiction.

To prove the second inequality we let $J \subset \mathcal{V}$ be an open subset and denote $J^* = \mathscr{S}(H)$. By Lemma 8.1 c) J^* is measurable, hence for any small $\delta > 0$ there is a closed set J^*_δ such that $J^*_\delta \subset J^*$ and $|J^*| - \delta \le |J^*_\delta| \le |J^*|$. This is possible because by Proposition 8.1 b) $\mathscr{S}$ is one-to-one modulo a set of measure zero. Let N_δ be an open set, $|N_\delta| < \delta$ containing the points where the inverse of $\mathscr{S}$ is not defined. We claim that there is k_0 such that

$$J^*_\delta \setminus N_\delta \subset J^*_k \stackrel{def}{=} \mathscr{S}_k(J), \quad \text{for any } k \ge k_0. \tag{8.5}$$

Here $\mathscr{S}_k$ is the mapping generated by u_k. Proof of (8.5) is by contradiction. If (8.5) fails then there is $x_k \in J^*_\delta \setminus N_\delta$ and $x_k \notin J^*_k$. We can assume that $x_k \to x_0$. Since $J^*_\delta \setminus N_\delta$ is closed it follows that $x_0 \in J^*_\delta \setminus N_\delta$. By definition of N_δ the inverse of $\mathscr{S}$ is one-to-one on $J^*_\delta \setminus N_\delta$. Thus there is a unique $Z_0 \in H$ such that $x_0 = \mathscr{S}(Z_0)$. Furthermore, there is an open neighborhood of Z_0 contained in J because J is open. If $H(x, \sigma_k, Z_k)$ is a supporting hyperboloid of u_k at x_k it follows from Lemma 7.2 that $x_k \in \mathscr{S}_k(Z_k), Z_k \to Z_0$. Thus for large k, $\{Z_k\}$ is in some neighborhood of $Z_0 \in J$ implying that $x_k \in J^*_k$ which contradicts our supposition. □

Proposition 8.2. *Let $f : \mathcal{U} \to \mathbb{R}$ and $g : \mathcal{V} \to \mathbb{R}$ be two nonnegative integrable functions. If $\mathcal{U} \subset \Pi$ and $\mathcal{V} \subset \Sigma$ are bounded domains such that the energy balance condition (1.3) and (1.18) hold then there exists a $B-$type weak solution to the problem (**RP**).*

Notice that we do not exclude the case $\mathcal{U} \cap \widehat{\mathcal{V}} \neq \emptyset$.

Proof. The proof of Proposition 8.2 is by approximation argument. Let $g_N = \sum_{i=1}^N C_i \delta_{Z_i}$ with $C_i \ge 0$ such that $\sum_{i=1}^N C_i = \int_{\mathcal{U}} f(x)dx, Z_i \in \Sigma$ and δ_{Z_i} are atomic measures supported at Z_i. For each g_N we construct a $B-$type solution u_N. Then sending $N \to \infty$ and using the compactness argument together with weak convergence of g_N to g, Lemma 8.1, one will arrive at desired result.

First, for each $Z \in \Sigma$ we define

$$\bar{a}(Z) = \frac{\varepsilon Z^{n+1} - \sqrt{(Z^{n+1})^2 + \rho^2}}{\varepsilon^2 - 1} \tag{8.6}$$

where

$$\rho(z) = \inf\{R > 0 : \mathcal{U} \subset B_z(R)\}. \tag{8.7}$$

Clearly $\bar{a}(Z)$ is the maximal value of larger semiaxis of hyperboloid H such that Γ_H is visible from $\mathcal{U}$ in the e_{n+1} direction. In other words $H(x, \bar{a}(Z), Z)$ is the lowest possible hyperboloid with focus $Z \in \Sigma$ such that

$H(x,\bar a(Z),Z)\ge 0$. Thus for $a\in(0,\bar a(Z)]$ we have $H(\cdot,a,Z)\in\mathbb{H}^+(\mathcal{U},\mathcal{V})$. To check (8.6) we fix Z and pick x_0 such that $\rho(z)=|x_0-z|$. Since the ratio of distances of x_0 from lower focus Z' and the plane $\Pi_d=\{X\in\mathbb{R}^{n+1}: X^{n+1}=Z^{n+1}-\bar a\varepsilon-\bar a/\varepsilon\}$ is ε, it follows that $|x_0Z'|=\varepsilon(Z^{n+1}-\bar a\varepsilon-\bar a/\varepsilon)$. On the other hand $|x_0Z|-|x_0Z'|=2\bar a$. Consequently, we find that $\sqrt{(Z^{n+1})^2+\rho^2(z)}=2\bar a+\varepsilon(Z^{n+1}-\bar a\varepsilon-\bar a/\varepsilon)$ which gives (8.6).

Next we define the maximal level $L_0=\sup\limits_{\mathcal{U}\times\mathcal{V}}H(x,\bar a(Z),Z)$. Since

$$\begin{aligned}\max_{x\in\mathcal{U}}H(x,\bar a(Z),Z) &= Z^{n+1}-\varepsilon\bar a(Z)-\bar a(Z)=\frac{\sqrt{(Z^{n+1})^2+\rho^2(z)}-Z^{n+1}}{\varepsilon-1}\\ &=\frac{\rho^2(z)}{(\varepsilon-1)(\sqrt{(Z^{n+1})^2+\rho^2(z)}+Z^{n+1})}\end{aligned}$$

it follows that

$$L_0=\sup_{\mathcal{V}}\frac{\rho^2(z)}{(\varepsilon-1)(\sqrt{(Z^{n+1})^2+\rho^2(z)}+Z^{n+1})}\le\frac{1}{\varepsilon-1}\sup_{\mathcal{V}}\frac{\rho^2(z)}{2Z^{n+1}}. \tag{8.8}$$

Next, we $H(\cdot,a,Z)$ by below for $a>0$ close to zero. By definition (7.3) we have that for this case $H(x,a,Z)\sim Z^{n+1}-\frac{\rho(z)}{\sqrt{\varepsilon^2-1}}$. We demand $Z^{n+1}-\frac{\rho(z)}{\sqrt{\varepsilon^2-1}}\ge 2L_0$ or equivalently in lieu of (8.8)

$$Z^{n+1}\ge\rho(z)\left[\frac{1}{\sqrt{\varepsilon^2-1}}+\frac{2\rho(z)}{(\varepsilon-1)(\sqrt{(Z^{n+1})^2+\rho^2(z)}+Z^{n+1})}\right].$$

But clearly $\frac{2\rho(z)}{(\varepsilon-1)(\sqrt{(Z^{n+1})^2+\rho^2(z)}+Z^{n+1})}\le 2/(\varepsilon-1)$. Therefore it is enough to assume that $Z^{n+1}\ge[\frac{2}{\varepsilon-1}+\frac{1}{\sqrt{\varepsilon^2-1}}]\rho(z)$ which is exactly (1.18). It follows that if Σ satisfies (1.18) then $\widetilde\Sigma=\Sigma+Me_{n+1}, M\gg 1$ also does.

Let $\mathbf{a}=(a_1,\dots,a_N), a_i\in(0,\bar a(Z_i)], i=1,\dots,N$ and set

$$H(\mathbf{a},x)=\min\big[H(x,a_1,Z_1),\dots,H(x,a_N,Z_N)\big].$$

We also let $\mathcal{E}_i(\mathbf{a})=\{x\in\mathcal{U}:H(\mathbf{a},x)=H(x,a_i,Z_i)\}$ be the i-th visibility sets and

$$\mathcal{A}^N=\left\{\mathbf{a}\in\prod_{i=1}^N(0,\bar a_i(Z_i)]:\int_{\mathcal{E}_i(\mathbf{a})}f\le C_i,\int_{\mathcal{E}_N(\mathbf{a})}f\ge C_N,\quad i=1,\dots,N-1.\right\}$$

From (1.18) it follows that $\mathcal{A}^N$ is not empty for taking $a_i, 1\le\le N-1$ close to zero and $a_N=\bar a_N(Z_N)$ one readily gets that such $\mathbf{a}$ is in $\mathcal{A}^N$.

The visibility sets $\mathcal{E}_i(\mathbf{a})$ enjoy the following property: if for some $a_k<\bar a(Z_k)$ we set $\mathbf{a}^k_\delta=(a_1,\dots,a_k+\delta,\dots,a_N)$ and $\mathbf{a}=(a_1,\dots,a_N)$ for $\delta>0$ small, then

$$\mathcal{E}_k(\mathbf{a})\subset\mathcal{E}_k(\mathbf{a}^k_\delta)\quad\text{while}\quad\mathcal{E}_i(\mathbf{a}^k_\delta)\subset\mathcal{E}_i(\mathbf{a}), i\ne k. \tag{8.9}$$

This can be seen for $N=2$ by simple geometric considerations, and general case is by induction.

Let $\mathfrak{a}=\sup\limits_{\mathbf{a}\in\mathcal{A}^N}\sum\limits_{i=1}^N a_i$ and $\hat{\mathbf{a}}\in\mathcal{A}^N$ be such that the supremum is realised, i.e. $\mathfrak{a}=\sum\limits_{i=1}^N\hat a_i$. We claim that $H(\hat{\mathbf{a}},x)$ solves the refractor problem with measure g_N. If not, then there is i_0, say $i_0=1$, such that $\int_{\mathcal{E}_1(\hat{\mathbf{a}})}f<C_1$. Then in view of the energy balance condition this implies $\int_{\mathcal{E}_N(\hat{\mathbf{a}})}f>C_N$. For $\delta>0$ small $F_N(\delta)=\int_{\mathcal{E}_N(\hat{\mathbf{a}}^1_\delta)}f(x)dx\ge C_N$ because $F_k(\delta)$ is continuous function of δ. Furthermore, using (8.9) it follows that $\mathbf{a}^1_\delta\in\mathcal{A}^N$ which is a contradiction. Now the proof of Theorem C1 follows from the above polyhedral approximation $H(\hat{\mathbf{a}},x)$ as $N\to\infty$ and the weak convergence of measures $\beta_{H,f}$, Lemma 8.1. □

9. An approximation lemma

9.1. **Refraction cone.** Recall that for smooth refractors the unit direction of the refracted ray is

$$Y=\varepsilon\left(e_{n+1}+\gamma\left[\sqrt{(\gamma\cdot e_{n+1})^2-\kappa}-\gamma\cdot e_{n+1}\right]\right),$$

see (3.4). This formula may be generalized for non smooth refractors as follows: let γ_1,γ_2 be the normals of two supporting planes of u at x. Then for any two constants c_1, c_2 the unit vector $\gamma_{c_1c_2}=\frac{c_1\gamma_1+c_2\gamma_2}{|c_1\gamma_1+c_2\gamma_2|}$ generates a mapping to the unit sphere $\mathbb{S}^{n+1}$ given by

$$\gamma_{c_1c_2}\mapsto\varepsilon\left(e_{n+1}+\gamma_{c_1c_2}\left[\sqrt{(\gamma_{c_1c_2}\cdot e_{n+1})^2-\kappa}-\gamma_{c_1c_2}\cdot e_{n+1}\right]\right).$$

Definition 9.1. *For $\gamma_1,\gamma_2\in\mathbb{S}^{n+1}$ the refractor cone at $\xi\in\mathbb{R}^{n+1}$ is defined as*

$$C_{\xi,\gamma_1,\gamma_2}=\left\{Z\in\mathbb{R}^{n+1}:\frac{Z-\xi}{|Z-\xi|}=\varepsilon\left(e_{n+1}+\gamma_{c_1c_2}\left[\sqrt{(\gamma_{c_1c_2}\cdot e_{n+1})^2-\kappa}-\gamma_{c_1c_2}\cdot e_{n+1}\right]\right)\right\}.$$

One can easily verify that $C_{\xi,\gamma_1,\gamma_2}$ is a convex cone. Indeed, for any $\gamma_0\perp\mathrm{Span}\{\gamma_1,\gamma_2\}$ we have that $\frac{Z-\xi}{|Z-\xi|}\cdot\gamma_0=\varepsilon(e_{n+1}\cdot\gamma_0)$. Thus $C_{\xi,\gamma_1,\gamma_2}$ is a cone.

In view of Lemma 7.1 $\|Du\|_\infty<\frac{1}{\sqrt{\varepsilon^2-1}}$ for any admissible $u\in\overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U},\mathcal{V})$, and $\sqrt{(\gamma\cdot e_{n+1})^2-\kappa}$ is well defined thanks to this gradient estimate.

9.2. **R-convexity of $\mathcal{V}$.**

Definition 9.2. *We say that $\mathcal{V}\subset\Sigma$ is $R-$convex with respect to a point $\xi\in[0,\infty)\times\mathcal{U}$ if for any two unit vectors γ_1,γ_2 the intersection $C_{\xi,\gamma_1,\gamma_2}\cap\mathcal{V}$ is connected. If $\mathcal{V}$ is $R-$convex with respect to any $\xi\in[0,\infty)\times\mathcal{U}$ then we simply say that $\mathcal{V}$ is $R-$convex.*

In particular a geodesic ball on the convex surface Σ is an example of $R-$convex $\mathcal{V}$.

9.3. **Local supporting function is also global.** In the Definition 7.1 of admissibility supporting hyperboloid H is saying above u in whole $\mathcal{U}$. Consequently, one may wonder if the locally admissible functions (i.e. H stays above u only in a vicinity of the contact point) are still in $\overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U},\mathcal{V})$. This issue was addressed by G. Loeper in [13] for the optimal transfer problems. We have

Lemma 9.1. *Under the condition (1.15) a local supporting hyperboloid is also global.*

The proof is very similar to that of in [13], [19] and hence omitted here. As an application of Lemma 9.1 we have the following approximation result.

Lemma 9.2. *If $u\in\overline{\mathbb{W}}_{\mathrm{H}}(B_r,\Sigma)$ then*

(i) *$u_\varepsilon(x)+K(r^2-|x|^2)\in\overline{\mathbb{W}}_{\mathrm{H}}(B_r,\widetilde{\Sigma})$ where u_ε is the standard mollification of u, $K>0$ and $\widetilde{\Sigma}=\Sigma+Me_{n+1}$ for some large constant $M>0$,*

(ii) *$u_\varepsilon(x)+K(r^2-|x|^2)$ is a classical subsolution of (5.2).*

Proof. (i) It is well known that u_ε is concave and $\|Du_\varepsilon\|_\infty\le\|Du\|_\infty<\frac{1}{\sqrt{\varepsilon^2-1}}$. Therefore if K is fixed then we can choose r so small that

$$\|D\bar{u}_\varepsilon\|_\infty\le\|Du\|_\infty+2Kr<\frac{1}{\sqrt{\varepsilon^2-1}}. \tag{9.1}$$

Moreover $K(r^2-|x|^2)$ is concave, hence $\bar u_\varepsilon = u_\varepsilon(x)+K(r^2-|x|^2)$ is concave too. Notice that $D^2\bar u_\varepsilon = D^2 u_\varepsilon - 2K\mathrm{Id}\le -2K\mathrm{Id}<0$ implying that $\bar u_\varepsilon$ is strictly concave. In order to bound the curvature of $\Gamma_{\bar u_\varepsilon}$ from below we recall that for fixed Z, $H(\cdot,a,Z)$ becomes flatter as $a\to\infty$ because

$$D^2H=-\frac{1}{(\varepsilon^2-1)\sqrt{a^2+\frac{|x-z|^2}{\varepsilon^2-1}}}\left[\mathrm{Id}-\frac{(x-z)\otimes(x-z)}{(\varepsilon^2-1)a^2+|x-z|^2}\right].$$

In particular, for large K and a we will have $-D^2\bar u_\varepsilon\ge 2K\mathrm{Id}\ge -D^2H$. Consequently, for each $x\in\mathcal U$ there is Z and $a>0$ such that $H(\cdot,a,Z)$ touches $\bar u_\varepsilon$ from above at x, in some neighbourhood of x. Furthermore, from Lemma 7.3 on confocal expansion we can choose $a,\widetilde Z$ so that $\widetilde Z\in\widetilde\Sigma=\Sigma+Me_{n+1}, M\gg 1$. Finally applying Lemma 9.1 we infer that $H(\cdot,a,\widetilde Z)$ is a global supporting hyperboloid of u at x and thus $\bar u_\varepsilon\in\overline{\mathbb W}_{\mathrm H}(\mathcal U,\widetilde\Sigma)$.

(ii) By direct computation we have

$$\begin{aligned}\mathcal M &= -D^2\bar u_\varepsilon-\frac{G(x,\bar u_\varepsilon,D\bar u_\varepsilon)}{\varepsilon\kappa}=-D^2u_\varepsilon+2K\mathrm{Id}-\frac{G(x,\bar u_\varepsilon,D\bar u_\varepsilon)}{\varepsilon\kappa}\ge\\ &\ge 2K\mathrm{Id}-\frac{G(x,\bar u_\varepsilon,D\bar u_\varepsilon)}{\varepsilon\kappa}.\end{aligned}\tag{9.2}$$

By definition, (5.1) we have

$$\frac{G(x,\bar u_\varepsilon,D\bar u_\varepsilon)}{\varepsilon\kappa}=\frac{[q+1](\mathrm{Id}-\varepsilon^2\kappa D\bar u_\varepsilon\otimes D\bar u_\varepsilon)}{\varepsilon\kappa t(x,\bar u_\varepsilon,D\bar u_\varepsilon)}\le\frac{C}{t(x,\bar u_\varepsilon,D\bar u_\varepsilon)}$$

with some tame constant $C>0$ depending only on ε. Recall that by (1.17) $t=\frac{(M+[Z^{n+1}-\bar u_\varepsilon])}{Y^{n+1}}\sim\frac{M}{c(\varepsilon)}, Z\in\Sigma$. Therefore choosing M large enough, one sees that $\mathcal M\ge\left[2K-\frac{Cc(\varepsilon)}{M}\right]\mathrm{Id}\ge K\mathrm{Id}$ if $K>\frac{Cc(\varepsilon)}{M}$. Fixing $K\ge\max[\frac{Cc(\varepsilon)}{M},\sup|h|^{\frac1n}]$, where h is defined by (5.4) and choosing r small enough such that (9.1) holds we finally arrive at $\det\left[-D^2\bar u_\varepsilon-\frac{G(x,\bar u_\varepsilon,D\bar u_\varepsilon)}{\varepsilon\kappa}\right]\ge|h|$ and the proof is complete. □

10. A-type weak solutions and the Legendre-like transform

In this section we are concerned with the second notion of weak solution to (**RP**). For $u\in\overline{\mathbb W}_{\mathrm H}(\mathcal U,\mathcal V)$ let us consider the mapping $\mathscr R_u:\mathcal U\to\Sigma$ defined as

$$\mathscr R_u(x)=\{Z\in\Sigma:\text{there is a supporting hyperboloid } H(\cdot,a,Z) \text{ of } u \text{ at } x\}.$$

Let $E\subset\mathcal U$ be a Borel set and put

$$\mathscr R_u(E)=\bigcup_{x\in E}\mathscr R_u(x).$$

Our primary goal is to prove that $\mathscr R_u(E)$ is measurable with respect to the restriction of $\mathcal H^n$ on Σ for any Borel set $E\subset\mathcal U$. That done, we can proceed as in [8] and establish that the set function $\alpha_{u,g}$ is σ-additive measure.

To take advantage of the geometric intuition coming from supporting hyperboloids of $u\in\overline{\mathbb W}_{\mathrm H}(\mathcal U,\mathcal V)$ it is convenient to define the Legendre-like transformation of u. Let $u\in\overline{\mathbb W}_{\mathrm H}(\mathcal U,\Sigma)$ and $Z\in\Sigma$ be a fixed point. Then the smallest semi-axis among all hyperboloids $H(\cdot,a,Z)$ that stay above u is

$$a_0=\sup_{a\in I(Z)}a,\quad I(Z)=\{a>0:H(x,a,Z)\ge u(x) \text{ in } \mathcal U\}.$$

Suppose that $H(\cdot,a_0,Z)$ touches u at $x_0\in\mathcal U$ then

$$u(x_0)=\psi(z)-a_0\varepsilon-\sqrt{a_0^2+\frac{(x_0-z)^2}{\varepsilon^2-1}}.$$

From here we can easily find that

$$a_0 = \frac{1}{\varepsilon^2 - 1}\left\{\varepsilon[u(x_0) - \psi(z)] - \sqrt{[u(x_0) - \psi(z)]^2 + (x_0 - z)^2}\right\}. \tag{10.1}$$

Alternatively, one can use the property that the distance of a point P on hyperboloid from lower focus Z' is ε times the distance of P from the hyperplane $\Pi_d = \{X \in \mathbb{R}^{n+1} : X^{n+1} = Z^{n+1} - a\varepsilon - \frac{a}{\varepsilon}\}$ (which in one dimensional case is the directrix). Since by definition of hyperboloid $|PZ| - |PZ'| = 2a$ and $|PZ'| = \varepsilon \mathrm{dist}(P, \Pi_d)$ we infer $|PZ| = 2a + \varepsilon([\psi(z) - u(x_0)] - a\varepsilon - \frac{a}{\varepsilon}) = -a(\varepsilon^2 - 1) + \varepsilon([\psi(z) - u(x_0)]$ and (10.1) follows.

10.1. **A-type weak solutions.**

Definition 10.1. *Let $u \in \overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \Sigma)$ then*

$$v(z) = \inf_{x\in\mathcal{U}}\left\{\varepsilon[\psi(z) - u(x)] - \sqrt{[u(x) - \psi(z)]^2 + (x - z)^2}\right\} \tag{10.2}$$

is called the Legendre-like transformation of u.

If $\mathrm{dist}(\mathcal{U}, \mathcal{V}) > 0$ and $\psi \in C^2$ then the function $\mathscr{L}_x(z) = \varepsilon[\psi(z) - u(x)] - \sqrt{[u(x) - \psi(z)]^2 + (x - z)^2}$ is C^2-smooth for any fixed $x \in \mathcal{U}$. Since v is the upper envelope of C^2 smooth functions $\mathscr{L}_x, x \in \mathcal{U}$ (x being the parameter) then $v(z)$ is semi-convex. Next lemma gives an important characterization of $v(z)$.

Lemma 10.1. *Let v be the Legendre-like transformation of $u \in \overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U}, \Sigma)$. Then*

(i) *$v(z) = \varepsilon[\psi(z) - u(x_0)] - \delta_u(x_0, z)$ if $Z = (z, \psi(z)) \in \mathscr{R}_u(x_0)$ where $\delta_u(x, z) = \sqrt{[u(x) - \psi(z)]^2 + (x - z)^2}$,*
(ii) *$v(z)$ is semi-concave.*

Proof. By definition $v(z)$ is locally bounded, non-negative, lower semi-continuous function. Let $\delta_u(x, z)$ denote the distance between the points of graph Γ_u and Σ. To check (i) we first observe that by definition of $v(z)$, see (10.2), we have $v(z) \le \varepsilon[\psi(z) - u(x_0)] - \delta_u(x_0, z)$. If $v(z) < \varepsilon[\psi(z) - u(x_0)] - \delta_u(x_0, z)$ it follows from (10.1) and the discussion above that $H(\cdot, a_0, Z)$ is a supporting hyperboloid of u at x_0, where $a_0 = (\varepsilon^2 - 1)^{-1}(\varepsilon[\psi(z) - u(x_0)] - \delta_u(x_0, z))$ because $Z \in \mathscr{R}_u(x_0) \subset \Sigma$. On the other hand, there is a sequence $\{x_k\}$ in $\mathcal{U}$ such that $x_k \to \bar{x}_0 \in \overline{\mathcal{U}}$ and $\lim_{x_k\to\bar{x}_0}(\varepsilon[\psi(z) - u(x_k)] - \delta_u(x_k, z)) = v(z)$. Setting $\bar{a}_0 = (\varepsilon^2 - 1)^{-1}v(z)$ we conclude that $H(\cdot, \bar{a}_0, Z)$ is touching Γ_u at $\bar{x}_0$. By construction $\bar{a}_0 < a_0$ and it follows from confocal expansion of hyperboloids 7.4 that $H(\cdot, \bar{a}_0, Z) > H(\cdot, a_0, Z)$ in $\mathcal{U}$. But this inequality is in contradiction with the fact that $H(\cdot, a_0, Z)$ is a supporting hyperboloid of u at x_0 and $H(\cdot, \bar{a}_0, Z)$ touches Γ_u at $\bar{x}_0$ whilst staying above Γ_u.

To prove (ii) we let $\mathscr{L}_{x_0}(y) = \varepsilon[\psi(z) - u(x_0)] - \delta_u(x_0, y)$. Then

$$v(y) = \inf_{x\in\mathcal{U}}\left\{\varepsilon[\psi(z) - u(x)] - \delta_u(x, y)\right\} \le \varepsilon[\psi(z) - u(x_0)] - \delta_u(x_0, y)$$

which implies that $v(y) \le \mathscr{L}_{x_0}(y)$ and $v(z) = \mathscr{L}_{x_0}(z)$, where $Z \in \mathscr{R}_u(x_0)$. We can regard $\mathscr{L}_{x_0}(y)$ as an upper supporting function of v at z. Differentiating $\mathscr{L}_{x_0}$ twice in z variable we see that $|D^2\mathscr{L}_{x_0}(z)| \le \frac{C}{(\mathrm{dist}(\mathcal{U},\mathcal{V}))^3}$ for some tame constant $C > 0$, consequently $v(z) - C|z|^2$ is concave for large $C > 0$. □

The main result of this section is contained in the following

Lemma 10.2. *Let $\mathcal{S} = \{Z \in \mathcal{V} :$ such that $Z \in \mathscr{R}_u(x_1) \cap \mathscr{R}_u(x_2), x_1 \ne x_2\}$. Then $\mathcal{S}$ has vanishing surface measure on Σ.*

Proof. Let us show that if $Z \in \mathcal{S}$ then the Legendre-like transformation of u is not differentiable at Z. This will suffice to conclude the proof because by definition v is semiconcave and hence by Aleksandrov's theorem twice differentiable almost everywhere. Let v be the Legendre-like transformation of u, then by Lemma 10.1 for any $Z\in\mathscr{R}_u(x_0)$ at which $v(z)$ is differentiable there holds

$$Dv(z)=\varepsilon D\psi(z)-(y(x)+D\psi(z)y^{n+1}(x)). \tag{10.3}$$

Indeed, $Dv(z)=\varepsilon D\psi(z)-\delta_u(x,z)^{-1}\left[(z-x)+D\psi(z)(\psi(z)-u(x))\right]$. From the definition of stretch function t it follows that $(z-x,\psi(z)-u(x))=Y\delta_u(x,z)$ where $Y=(y,y^{n+1})$ is the unit direction of the refracted ray and (10.3) follows. Consequently, if $x_1\neq x_2$ such that $\mathscr{R}_u(x_1)\cap\mathscr{R}_u(x_2)\ni Z$ then we must have

$$Dv(z)=-\frac{z-x_i+D\psi(z)(\psi(z)-x_i)}{\delta_u(x_i,z)}+\varepsilon D\psi(z),\quad i=1,2.$$

Equating the right hand sides for $i=1$ and $i=2$ we obtain

$$\frac{z-x_1+D\psi(z)(\psi(z)-x_1)}{\delta_u(x_1,z)}=\frac{z-x_2+D\psi(z)(\psi(z)-x_2)}{\delta_u(x_2,z)}$$

With the aid of this observation and (10.3) we can rewrite the last line as follows

$$y_1+D\psi(z)y_1^{n+1}=y_2+D\psi(z)y_2^{n+1}\text{ in }\mathbb{R}^n\quad\Rightarrow\quad Y_1+(D\psi(z),-1)y_1^{n+1}=Y_2+(D\psi(z),-1)y_2^{n+1},\text{ in }\mathbb{R}^{n+1}.$$

The last identity implies that Y_1-Y_2 is collinear to the normal of Σ at Z. Consequently, from the assumption (1.11) (see also (1.17)) we obtain that this is possible if and only if $Y_1=Y_2$. Next, from $Y_1=Y_2$ we have $y_1=y_2$ and consequently we conclude that

$$\frac{z-x_1}{\delta_u(x_1,z)}=\frac{z-x_2}{\delta_u(x_2,z)}. \tag{10.4}$$

Taking the reciprocal of both sides in the last identity and recalling the definition of the distance $\delta_u(x,z)$ one gets

$$\frac{u(x_1)-\psi(z)}{|x_1-z|}=\frac{u(x_2)-\psi(z)}{|x_2-z|}$$

yielding

$$\begin{aligned}u(x_1)&=\psi(z)+\frac{|z-x_1|}{|z-x_2|}(u(x_2)-\psi(z))\\&=\psi(z)+\frac{\delta_u(x_1,z)}{\delta_u(x_2,z)}(u(x_2)-\psi(z)).\end{aligned}$$

On the other hand $Y_1^{n+1}=Y_2^{n+1}$ gives $u(x_1)-u(x_2)=\delta_u(x_2,z)-\delta_u(x_1,z)$ and hence combining this with the last equation yields

$$\psi(z)\left[1-\frac{\delta_u(x_1,z)}{\delta_u(x_2,z)}\right]-u(x_2)\left[1-\frac{\delta_u(x_1,z)}{\delta_u(x_2,z)}\right]=\delta_u(x_1,z)-\delta_u(x_2,z).$$

If $\delta_u(x_2,z)\neq\delta_u(x_1,z)$ then the last equality implies $u(x_2)-\psi(z)=\sqrt{(u(x_2)-\psi(z))^2+(z-x_2)^2}$. Hence $x_2=z$ and by (10.4) $x_1=x_2$, which is contradiction. Thus we must have $\delta_u(x_2,z)=\delta_u(x_1,z)$ and in view of (10.4) this implies that $x_1=x_2$, again contradicting our supposition. Therefore we infer that $v(z)$ cannot be differentiable at z. By Rademacher's theorem $v(z)$ is differentiable a.e. in z. Thus $\mathcal{S}$ has vanishing surface measure. □

Corollary 10.1. *For any $u\in\overline{\mathbb{W}}_{\mathrm{H}}(\mathcal{U},\mathcal{V})$ and any Borel subset $E\subset\mathcal{U}$ the set function*

$$\alpha_{u,g}(E)=\int_{\mathscr{R}_u(E)}gd\mathcal{H}^n \tag{10.5}$$

is a Radon measure.

Proof. In order to show that $\alpha_{u,g}$ is Radon measure it suffices to check that $\widetilde{\mathscr{F}}=\{E\subset\mathcal{U}:\mathscr{R}_u(E)\text{ is measurable}\}$ is a $\sigma-$algebra. This can be done exactly in the same way as in the proof of Proposition 8.1 c). It remains to recall that by Lemma 7.2, $\widetilde{\mathscr{F}}$ contains the closed sets. □

Definition 10.2. *A function $u \in \overline{\mathbb{W}}_{\rm H}(\mathcal{U}, \mathcal{V})$ is said to be A-type weak solution of (***RP***) if $\int_E f(x)dx = \alpha_{u,g}(E)$ or any Borel set $E \subset \mathcal{U}$ and*

$$\overline{\mathcal{V}} \subset \overline{\mathscr{R}_u(\mathcal{U})}, \qquad |\{x \in \mathcal{U} : \mathscr{R}_u(x) \not\subset \mathcal{V}\}| = 0 \tag{10.6}$$

This definition is natural, stating that the target domain $\mathcal{V}$ is covered by the refracted rays and the endpoints of those rays that after refraction do not strike $\mathcal{V}$ constitute a null set on $\mathcal{U}$. We shall establish the existence of A-type weak solution in the next section.

In closing this section we state the weak convergence result for the α-measures, see Corollary 10.1.

Lemma 10.3. *Let u_k be a sequence of A-type weak solutions and α_k is the corresponding measure, defined by (10.5). If $u_k \to u$ uniformly on compact subsets of $\mathcal{U}$ then u is $R-$concave and α_k weakly converges to $\alpha_{u,g}$.*

The proof is very similar to that of Lemma 8.1 (modulo minor adjustments) and hence omitted.

11. Comparing A and B type weak solutions: Proof of Theorem C3-4

Let $\varphi : \mathbb{R}^N \to \mathbb{R}^n$ be a Borel mapping and $\mu(\mathbb{R}^N) = \nu(\mathbb{R}^n) < \infty$ with μ, ν being two Radon measure on $\mathbb{R}^N$ and $\mathbb{R}^n$, respectively. Then φ induces a (push-forward) measure on $\mathbb{R}^n$ defined by $\varphi_\#\mu(E) = \mu(\varphi^{-1}(E))$ for Borel subsets $E \subset \mathbb{R}^n$. We say that a Borel mapping φ *measure preserving* if

$$\varphi_\#\mu(E) = \nu(E) \qquad \text{for any Borel set } E \subset \mathbb{R}^n. \tag{11.1}$$

By the change of variables formula (11.1) can be rewritten in the following equivalent form

$$\int h(\varphi(x))d\mu = \int h(y)d\nu, \qquad \forall h \in C(\mathbb{R}^n), \tag{11.2}$$

see [3].

Remark 11.1. *If $u \in \mathbb{W}^+(\mathcal{U}, \Sigma)$ is the B-type solution of (***RP***), the existence of which is established in Section 8, then taking $\varphi(Z) = \mathscr{S}_u(Z), N = n+1$, $d\mu = gd\mathcal{H}^n$, and ν being the Lebesgue measure one immediately observes that $\mathscr{S}_u$ is measure preserving in the sense of (11.1) or (11.2).*

Lemma 11.1. *If $\mathscr{R}_u(x) \subset \mathcal{V}$ for a.e. $x \in \mathcal{U}$ then $\mathscr{R}_u(E) \subset \mathrm{Hull}(\mathcal{V})$, where $\mathrm{Hull}(\mathcal{V})$ is the R-convex hull of $\mathcal{V}$ defined as the smallest R-convex subset of Σ containing $\mathcal{V}$.*

Proof. We only have to consider the points where u is non-differentiable. Let u be non-differentiable at $x_0 \in \mathcal{U}$ and suppose that γ_1, γ_2 are the normals of two supporting planes of u at x_0. The ray with endpoint x_0 after reflection will lie in the reflector cone $\mathcal{C}_{\xi_0,\gamma_1,\gamma_2}$, with $\xi_0 = (x_0, u(x_0))$ and the reflected ray will strike $\mathrm{Hull}(\mathcal{V})$, because $\mathcal{C}_{\xi_0,\gamma_1,\gamma_2} \cap \mathrm{Hull}(\mathcal{V})$ is connected. Considering all normals of supporting planes at x_0 we obtain the desired result. □

Proposition 11.1. *Let Σ be $R-$convex with respect to $Q_m = \mathcal{U} \times (0, m), m > 0$ and the densities f, g are positive. Then B-type weak solution is also of A-type.*

Proof. We split the proof into three parts.

1) First we show that for any compact $K_1 \subset \mathcal{U}$ there holds $\int_{K_2} gd\mathcal{H}^n \geq \int_{K_1} f(x)dx$ with $K_2 = \mathscr{R}_u(K_1)$. In other words the B-type solution is A-type subsolution. It is worthwhile to point out that for the proof of this inequality we don't need $\mathcal{V}$ to be $R-$convex. Take $\eta \in C(\Sigma)$ such that $\eta \equiv 1$ on $K_2 \subset \Sigma$ and $0 \leq \eta \leq 1$. From (11.2) we see that

$$\int_{\mathcal{V}} \eta gd\mathcal{H}^n = \int_{\mathcal{U}} \eta(\mathscr{R}_u(x))f(x)dx \geq \int_{K_1} f(x)dx.$$

Letting η to decrease to the characteristic function of K_2, $h \downarrow \chi_{K_2}$ we infer

$$\int_{K_2} g d\mathcal{H}^n \geq \int_{K_1} f(x)dx. \tag{11.3}$$

Notice by Corollary 10.1 the measure $\alpha_{u,g}$ is Borel regular, therefore in the last inequality K_1 can be replaced by any Borel subset of $\mathcal{U}$. As a result we conclude from (11.3) that

$$\text{if } \mathcal{H}^n(\mathscr{R}_u(E)) = 0 \text{ then } |E| = 0. \tag{11.4}$$

2) Next, we prove the converse estimate of (11.3). Here we will utilize the $R-$convexity of $\mathcal{V}$. Take any compact $K_1 \in \mathcal{U}$ and apply Lemma 10.2 to conclude $\mathcal{H}^n\big(\mathscr{R}_u(K_1) \cap \mathscr{R}_u(\mathcal{U} \setminus K_1)\big) = 0$. Let us show that

$$|\mathscr{R}_u^{-1}(\mathscr{R}_u(K_1)) \setminus K_1| = 0 \tag{11.5}$$

where $\mathscr{R}_u^{-1}(\mathscr{R}_u(K_1))$ is the pre-image of $\mathscr{R}_u(K_1)$. Denote $E = \mathscr{R}_u^{-1}(\mathscr{R}_u(K_1))$ and $G = K_1$. If $\mathcal{H}^n(E \setminus G) = 0$ then in view of (11.4) we obtain (11.5). Indeed, form the identity (8.2) it follows that

$$\begin{aligned} |\mathscr{R}_u(E \setminus G)| &= \Big|[\mathscr{R}_u(E) \setminus \mathscr{R}_u(G)] \bigcup [\mathscr{R}_u(E \setminus G) \cap \mathscr{R}_u(G)]\Big| \\ &= |\mathscr{R}_u(E \setminus G) \cap \mathscr{R}_u(G)| \\ &= 0 \end{aligned} \tag{11.6}$$

where to get the last line we used the definitions of E and G in order to obtain $\mathscr{R}_u(E)\setminus\mathscr{R}_u(G) = \mathscr{R}_u(K_1)\setminus\mathscr{R}_u(K_1) = \emptyset$ and Lemma 10.2. Thus (11.4) implies $0 = |E \setminus G| = |\mathscr{R}_u^{-1}(\mathscr{R}_u(K_1)) \setminus K_1|$.

Let $h \in C(\Sigma)$ such that $0 \leq h \leq 1$ and $h \geq \chi_{\mathscr{R}_u(K_1)}$. Since $\mathcal{V}$ is $R-$convex it follows that $\mathscr{R}_u(K_1) \subset \text{Hull}\mathcal{V}$, see Lemma 11.1. If u is a B-type weak solution then (11.2) holds, see Remark 11.1. Therefore

$$\begin{aligned} \int_{\mathcal{U}} \eta(\mathscr{R}_u(x)) f(x)dx &= \int_{\mathcal{V}} \eta g d\mathcal{H}^n \\ &= \int_{\text{Hull}(\mathcal{V})} \eta g d\mathcal{H}^n \\ &\geq \int_{\mathscr{R}_u(K_1)} g d\mathcal{H}^n. \end{aligned}$$

Letting $\eta \to 0$ on compact subsets of $\mathcal{V} \setminus \mathscr{R}_u(K_1)$, it follows that $\eta(\mathscr{R}_u(x))$ uniformly converges to zero one the compact subsets of $\mathcal{U} \setminus \mathscr{R}_u^{-1}(\mathscr{R}_u(K_1))$. Consequently

$$\int_{\mathscr{R}_u(K_1)} g d\mathcal{H}^n \leq \int_{\mathcal{U}} \eta(\mathscr{R}_u(x)) f(x)dx \longrightarrow \int_{\mathscr{R}^{-1}(\mathscr{R}_u(K_1))} f(x)dx = \int_{K_1} f(x)dx$$

where the last line follows from (11.5). This implies that u is a supersolution.

3) It remains to check that u verifies the boundary condition (10.6). Suppose that there is $Z_0 \in \overline{\mathcal{V}}$ such that $Z_0 \notin \overline{\mathscr{R}_u(\mathcal{U})}$. Since u is of B-type, it follows that $\mathscr{S}_u(\mathcal{V}) = \mathcal{U}$ implying $x_0 \in \mathscr{S}_u(Z_0)$ in other words, there is a supporting hyperboloid $H(x, a_0, Z_0)$ at x_0. Thus $Z_0 \in \mathscr{R}_u(x_0)$which yields $\overline{\mathcal{V}} \subset \overline{\mathscr{R}_u(\mathcal{U})}$. From energy balance condition we have

$$\int_{\overline{\mathscr{R}_u(\mathcal{U})}} g d\mathcal{H}^n = \int_{\mathcal{U}} f(x)dx = \int_{\mathcal{V}} g d\mathcal{H}^n \quad \Rightarrow \quad \int_{\overline{\mathscr{R}_u(\mathcal{U})}\setminus\mathcal{V}} g = 0.$$

This yields $|\{x \in \mathcal{U} : \mathscr{R}_u(x) \not\subset \mathcal{V}\}| = 0$ for $f, g > 0$. □

Remark 11.2. *We always have $\overline{\mathcal{V}} \subset \mathscr{R}_u(\mathcal{U})$, however if in addition Σ is $R-$convex then it follows that $\mathscr{R}_u(\mathcal{U}) \subset \mathcal{V}$. Thus we get the equality $\overline{\mathscr{R}_u(\mathcal{U})} = \overline{\mathcal{V}}$ for R-convex $\mathcal{V}$.*

11.1. **Existence of A-type weak solutions: Proof of Theorem C4.** Suppose that $\mathcal{V}\subset\Sigma$ and let Hull($\mathcal{V}$) be the $R-$convex hull of $\mathcal{V}$. For small $\delta,\delta'>0$ we consider

$$g_\delta(Z)=\begin{cases} g(Z)-\delta & \text{if } Z\in\mathcal{V}\\ \delta' & \text{if } Z\in \text{Hull}(\mathcal{V})\setminus\mathcal{V}\end{cases} \tag{11.7}$$

where we choose δ,δ' so that g_δ satisfies the energy balance condition (1.3). By Proposition 8.2 for each g_δ there is a B-type weak solution which according to Proposition 11.1 is also of A-type. Moreover, from Remark 11.2 we infer

$$\overline{\mathscr{R}_{u_\delta}(\mathcal{U})}=\overline{\mathcal{V}}. \tag{11.8}$$

Sending $\delta\to0$ we obtain from Lemma 10.3 that $u_\delta\to u$ and u is an A-type solution, i.e. (10.5) is satisfied, and

$$\overline{\mathcal{V}}\subset\overline{\mathscr{R}_u(\mathcal{U})}. \tag{11.9}$$

Since u is second order differentiable a.e. in $\mathcal{U}$ it follows that $\mathscr{R}_u$ is defined for a.e. $x\in\mathcal{U}$. Finally we want to show that $|S|=0$ where $S=\{x\in\mathcal{U}:\exists Z\in\mathscr{R}_u(x) \text{ such that } Z\in\mathscr{R}_u(\mathcal{U})\setminus\mathcal{V}\}$. Indeed, from energy balance condition (1.3) we have

$$\begin{aligned}\int_S f(x)dx &= \int_{\mathcal{U}} f(x)dx-\int_{\mathcal{U}\setminus S} f(x)dx=\\ &=\int_{\mathcal{U}} f(x)dx-\int_{\mathcal{V}} gd\mathcal{H}^n=0.\end{aligned}$$

Since $f>0$ we conclude that $|S|=0$ and hence (10.6) holds and u is a weak A-type weak solution of (**RP**). □

Remark 11.3. *As the proof of Proposition 11.1 exhibited if $\mathcal{V}$ is $R-$convex then $S=\emptyset$. If $S\neq\emptyset$ then u is only Lipschitz continuous. Therefore if $\mathcal{V}$ is not R-convex then u may not be C^1 smooth, see Introduction. It is worthwhile to point out that even if $S=\emptyset$ then u may not be C^1, and hence further assumptions must be imposed to assure the smoothness of u.*

12. Dirichlet's problem

This section concerns the Dirichlet problem for A-type weak solutions. We formally rewrite the equation (5.2) below

$$\mathcal{F}[u](x)=\frac{f(x)}{g\circ\mathscr{R}_u(x)},\qquad x\in\mathcal{U}, \tag{12.1}$$

where for $u\in C^2(\mathcal{U})$, $\mathcal{F}[u](x)$ is the determinant of the Jacobian matrix of $\mathscr{R}_u$. For non smooth solutions we give the following definition

Definition 12.1. *A function $u\in\overline{\mathbb{W}}_{\rm H}(\mathcal{U},\Sigma)$ is said to be a weak A-subsolution of (12.1) if for any Borel set E*

$$\int_{\mathscr{R}_u(E)} gd\mathcal{H}^n\geq\int_E f(x)dx. \tag{12.2}$$

If $\alpha_{u,g}(E)=\int_E f(x)dx$ then we say that u is a weak A-solution. The class of all generalized A-subsolutions is denoted by $\mathcal{AS}^+(\mathcal{U})$.

For smooth and bounded $D \subset \Sigma$ and smooth function $\varphi : \overline{D} \to \mathbb{R}$ let us consdier the Dirichlet problem

$$\begin{cases} \mathcal{F}[u](x) = \dfrac{f(x)}{g \circ \mathscr{R}_u(x)}, & x \in D, \\ u = \varphi & x \in \partial D. \end{cases} \tag{12.3}$$

Our main objective here is to prove the existence and uniqueness of A-type weak solution to (12.3) for a smooth boundary data. In fact, for our purposes it suffices to consider the case where D is a ball of small radius. At this point we first we establish the following comparison principle.

Proposition 12.1. *Let u_i be weak solutions of (12.1) in $\mathcal{U}$ with $f = f_i, i = 1, 2$, where $\Omega \subset \Pi$ is a smooth, bounded domain. Suppose that $\mathscr{R}_{u_1}(\Omega) \subset \overline{\Sigma}$, $f_1 < f_2$ in Ω and $u_1 \le u_2$ on $\partial\Omega$. If Γ_1, the graph of u_1, lies in the region $\mathcal{D}$ then we have $u_1 \le u_2$ in Ω.*

Proof. Suppose that $\Omega_1 = \{x \in \Omega : u_1(x) > u_2(x)\}$ is not empty. Let $x_0 \in \Omega_1$ and $H(x, a_0, Z_0), Z_0 \in \Sigma$ is a supporting hyperboloid of u_2 at x_0. From the confocal expansion of hyperboloids (see subsection 7.4) we infer that $H(x, a_0+s, Z_0)$ is a supporting hyperboloid of u_1 at an interior point $x_1 \in \Omega_1$ for some $s > 0$. Thus $H(x, a_0+s, Z_0)$ is a local supporting hyperboloid of u_1. Since Γ_{u_1} is in the regularity domain $\mathcal{D}$, where (1.11)-(1.15) are fulfilled, we can apply Lemma 9.1 to conclude that $H(x, a_0+s, Z_0)$ is also a global supporting hyperboloid of u_1. Therefore

$$\mathscr{R}_{u_2}(\Omega_1) \subset \mathscr{R}_{u_1}(\Omega_1)$$

implying

$$\int_{\Omega_1} f_1 dx < \int_{\Omega_1} f_2 dx = \int_{\mathscr{R}_{u_2}(\Omega_1)} g d\mathcal{H}^n \le \int_{\mathscr{R}_{u_1}(\Omega_1)} g d\mathcal{H}^n = \int_{\Omega_1} f_1 dx$$

which gives a contradiction. Thus $\Omega_1 = \emptyset$. □

12.1. **Discrete Dirichlet problem.** To outline our next two steps, we note that for the classical Monge-Ampère equation the standard way of proving the existence of globally smooth solutions to Dirichlet's problem with, say, $\varphi \in C^4(\overline{D})$ is to employ the continuity method combined with standard mollification argument, see [16]. Moreover, in this argument φ must be a subsolution. In order to tailor a similar proof for (12.3) we will mollify our weak A-solution, add $K(r^2 - |x - x_0|^2), K \gg 1$ and consider its restriction to $B_r(x_0) \subset \mathcal{U}$, a ball with sufficiently small radius. Such function turns out to be classical subsolution if K some large and small $r > 0$. Consequently, from continuity method one can obtain existence of a smooth solution to Dirichlet's problem in $B_r(x_0)$. Finally employing the known C^2 a priori estimates and comparison principle, Proposition 12.1, the proof of Theorem D will follow, see Section 13 for more details.. Our approach most closely follows that proposed by Xu-Jia Wang [21].

Let $\{b_i\} \subset \partial D$ be a sequence of points on the boundary of D and $\{a_i\} \subset D$. Let $A_N = \{a_1, \dots, a_N\}$ and $B_N = \{b_1, \dots, b_N\} \subset \partial D$, for each fixed $N \in \mathbb{N}$. Furthermore, let $\nu_k(x)$ be atomic measures supported at $a_k, 1 \le k \le N$ and let

$$\mathcal{F}[v](x) = \nu_k(x) \frac{f(x)}{g \circ \mathscr{R}_v(x)}. \tag{12.4}$$

Proposition 12.2. *Let $\underline{u} \in \overline{\mathbb{W}}^0_{\mathrm{H}}(D, \Sigma)$ be a polyhedral subsolution of (12.4), i.e. $\mathcal{F}[\underline{u}](x) \ge \nu_k(x) \frac{f(x)}{g \circ \mathscr{R}_{\underline{u}}(x)}$ at $a_k \in A_N$. Then there is a unique A-type weak solution to (12.4) verifying the boundary condition $u = \underline{u}$ on B_N.*

Proof. We want to construct a sequence of subsolutions $\{u_m\}_{m=0}^{\infty}$ converging to the solution of discrete problem. Set $u_0 = \underline{u}$ and define u_1 such that $u_1 \le u_0$ in A_N, $u_1(b_i) = u_0(b_i), b_i \in B_N$ and $\alpha_{u_1,g}(a_i) \le \alpha_{u_0,g}(a_i)$ for $a_i \in A_N$. It is convenient to introduce the class of hyperboloids

$$\Phi_{0,\delta}(a_1) = \left\{ P \in \mathbb{H}^+(D,\Sigma) : \begin{array}{c} H(a_i) \ge u_0(a_i), i \ne 1, \\ H(a_1) \ge u_0(a_1) - \delta, \\ H(b_j) \ge u_0(b_j), 1 \le j \le N \end{array} \right\}$$

for $\delta > 0$ and let

$$T_1^\delta u_0 = \inf_{H \in \Phi_{0,\delta}(a_1)} H.$$

Let $\delta_1 > 0$ be the largest δ for which $T_1^{\delta_1} u_0$ is a subsolution to (12.4) on A_N. Consequently, by setting $u_{0,1} = T_1^{\delta_1}$ we can proceed by induction and define the k-th class

$$\Phi_{0,\delta}(a_k) = \left\{ H \in \mathbb{H}^+(D,\Sigma) : \begin{array}{c} H(a_i) \ge u_{0,k-1}(a_i), i \ne k, \\ H(a_k) \ge u_{0,,k-1}(a_k) - \delta, \\ H(b_j) \ge u_{0,,k-1}(b_j), 1 \le j \le N \end{array} \right\}$$

and take $T_k^\delta u_0 = \inf\limits_{H \in \Phi_{0,\delta}(a_k)} H$. Therefore, one can successively construct the functions $u_{0,k} = T_k^{\delta_k} u_{0,k-1}$ where $\delta_k > 0$ is the largest number for which $T_k^\delta u_{0,k-1}$ is a subsolution to (12.4) in A_N. Taking the second subsolution in the approximating sequence to be $u_2(x) \overset{def}{=} T_N^{\delta_N} u_{0,N-1}$ we get, by construction, that $\alpha_{u_0,g}(a_i) \le \alpha_{u_1,g}(a_i)$, since we have the inclusions $\Phi_{l,\delta}(a_k) \subset \Phi_{l+1,\delta}(a_k)$ at a_k as we proceed. Therefore we have a sequence of solutions u_m to the Dirichlet problem in A_N such that

$$\alpha_{u_m,g}(a_i) \le \alpha_{u_{m-1},g}(a_i),$$
$$u_m(a_i) \le u_{m-1}(a_i),$$
$$u_m(b_i) = u_{m-1}(b_i).$$

The first two inequalities are obvious. As for the boundary condition we note that $u_0(b_i) \le u_1(b_i)$ by construction. If $u_0(b_i) < u_1(b_i)$ then by taking $\min[H_i(x), u_1(x)]$, where $H_i(x) \in \mathbb{H}^+(D,\Sigma)$ is a supporting hyperboloid of u_0 at b_i we see that $\min[H_i(x), u_1(x)]$ belongs to the corresponding Φ class. Thus $u_0(b_i) = u_1(b_i)$.

From Lemma 7.2 we conclude that $u \in \overline{\mathbb{W}}_{\mathrm{H}}(D,\Sigma)$ and in view of Lemma 10.3 $\alpha_{u_m,g} \rightharpoonup \alpha_{u,g}$ weakly. Thus $u = \lim\limits_{m\to\infty} u_m$ is a solution to the discrete problem in A_N with $u(b_i) = \underline{u}(b_i), b_i \in B_N$. □

12.2. **General case.** Perron's method, used in the proof of above proposition, can be strengthened in order to establish the solvability of the general Dirichlet problem. To do so we take $\{a_i\}_{i=1}^\infty \subset D$ and $\{b_i\}_{i=1}^\infty \subset \partial D$ to be dense subsets and $A_N = \{a_1, \dots, a_N\} \subset D, B_N = \{b_1, \dots, b_N\} \subset \partial D$.

Proposition 12.3. *Let $\underline{u} \in \mathcal{AS}^+(D,\Sigma)$. Then there exists a unique weak solution u to the Dirichlet problem*

$$\begin{cases} \mathcal{F}[u] = \dfrac{f(x)}{g \circ \mathscr{R}_u(x)} & \text{in } D, \\ u(x) = \underline{u}(x) & \text{on } \partial D. \end{cases} \tag{12.5}$$

Proof. For $\delta > 0$ we denote $D_\delta = \{x \in D : \operatorname{dist}(x, \partial D) > \delta\}$ and take $\eta(x)$ to be a smooth function such that $0 \le \eta(x) \le 1$, $\eta \equiv 1$ in $D_{2\delta}$ and $\eta \equiv 0$ in $D \setminus D_\delta$. Consider the equation

$$\mathcal{F}[v](x) = \nu_k(x) J(v(x)) \eta_\delta(x) \frac{f(x) - \delta}{g \circ \mathscr{R}_v(x)} \tag{12.6}$$

where $\nu_k(x)$ is a positive measure supported at $a_k \in A_N$ and

$$J(t) = \begin{cases} 1 & \text{if} \quad 0 \le t \le \sup_D \underline{u}, \\ \frac{2\sup_D \underline{u} - t}{\sup_D \underline{u}} & \text{if} \quad \sup_D \underline{u} \le t \le 2\sup_D \underline{u}, \\ 0 & \text{if} \quad t > 2\sup_D \underline{u}. \end{cases} \tag{12.7}$$

Consider the class

$$\mathbb{W}^+_{N,\underline{u}} = \left\{ v \in \overline{\mathbb{W}}^0_{\mathrm{H}}(D,\Sigma) : \mathcal{F}[v] \ge \nu_k J(v)\eta_\delta \frac{f-\delta}{g\circ \mathscr{R}_u} \text{ and } v \ge \underline{u} \text{ on } B_N \right\}. \tag{12.8}$$

Clearly $\mathbb{W}^+_{N,\underline{u}}$ is not empty since $H(\cdot, a, Z)$ is in this class if $a > 0$ is sufficiently small. We claim that if $v_{N,\delta} = \inf_{\mathbb{W}^+_{N,\underline{u}}} v$ then $v_{N,\delta}$ solves (12.1) in the sense of Definition 12.1 and $v_{N,\delta}(b_i) = \underline{u}(b_i), b_i \in B_N$.

It is easy to see that $\alpha_{v_{N,\delta},g}(a_k) = v_k(a_k)J(v_{N,\delta})\eta_\delta(a_k)\left(f(a_k) - \delta\right)$. Indeed, if $v_{N,\delta}$ is a strict subsolution at a_i, i.e. for some a_i we have $\alpha_{v_{N,\delta},g}(a_i) > v_k(a_i)J(v_{N,\delta})\eta_\delta(a_i)(f(a_i) - \delta)$, then we can push $\Gamma_{v_{N,\delta}}$ downward by some $\delta > 0$, decreasing the α measure at a_i, which, however, will be in contradiction with the definition of $v_{N,\delta}$. Thus $v_{N,\delta}$ is a solution of the equation (12.6).

Next, we check the boundary condition. Choose $H_i \in \mathbb{H}^+(\mathcal{U},\Sigma)$ such that $H_i > v_\delta$ in $\mathcal{U}_\delta$ and passes through $(b_i, \underline{u}(b_i))$. Such H_i exists because by construction $v_{N,\delta}(a_i) \le \underline{u}(a_i)$ and $\delta > 0$.

For $\widetilde{H}_i = \min[H_i, v_{N,\delta}]$, by construction, we see that $\mathcal{F}[\widetilde{H}_i] \ge \nu_k J(\widetilde{H}_i)\eta_\delta \frac{f-\delta}{g\circ\mathscr{R}_{\widetilde{H}_i}}$ at a_i. Thus $\widetilde{H}_i \in \mathbb{W}^+_{N,\underline{u}}$. Hence

$$v_{N,\delta}(b_i) = \inf_{H\in\mathbb{W}^+_{N,\underline{u}}} H(b_i) \le \widetilde{H}_i(b_i) = \underline{u}(b_i).$$

Now the desired solution can be obtained via a standard compactness argument that utilizes the estimates of Lemma 7.1 and Lemma 10.3. More precisely, for fixed $\delta > 0$ we send $N \to \infty$ and obtain a function v_δ that solves the equation $\mathcal{F}[v_\delta] = J(v_\delta)\eta_\delta \frac{f-\delta}{g\circ\mathscr{R}_{v_\delta}}$. To show that $v_\delta = \underline{u}$ on ∂D we take $x_0 \in \partial D$ and again use the comparison with $\min[H_0, v_\delta]$ for a suitable $H_0 \in \mathbb{H}^+(\mathcal{U},\Sigma)$ such that $H_0(x_0) = \underline{u}(x_0)$. Thus, from Proposition 12.1 we conclude that $v_\delta \le \underline{u}$ in D. Finally sending $\delta \downarrow 0$ and employing the estimate of Lemmas 7.1 and 10.3 we arrive at desired result. □

13. Proof of Theorem D

To fix the ideas we assume that $x_0 = 0 \in \mathcal{U}$ and $B_r = B_r(0) \subset \mathcal{U}$. Let $u^\pm_{s,\delta}$ be the solutions to

$$\begin{cases} \mathcal{F}[u^\pm_{s,\delta}] = \frac{f\pm\delta}{\eta g\circ Z_{u^\pm_{s,\delta}}} & \text{in } B_r \\ u^\pm_{s,\delta} = \widetilde{u}_s & \text{on } \partial B_r \end{cases} \tag{13.1}$$

where $\widetilde{u}_s = u_s + K(r^2 - |x|^2)$, $K > 0$ and u_s is a mollification of the weak solution u. By Lemma 9.2 $\widetilde{u}_s$ is a subsolution (for appropriate choice of constants K and r) and hence by Proposition 12.3 the weak solution to Dirichlet's problem exists. Note that for the Dirichlet problem we have to consider the modified receiver $\widetilde{\Sigma}$ to guarantee that $\widetilde{u}_s$ is admissible, see Lemma 9.2. In order to show the existence of smooth solutions we apply the continuity method: Let $\underline{w} \in \mathcal{AS}^+(B_r, \widetilde{\Sigma}) \cap C^\infty(B_r)$ and for $t \in [0,1]$ consider the solutions to the Dirichlet problem

$$\begin{cases} \mathcal{F}[w^t] = t\frac{f}{hg\circ Z_w} + (1-t)\mathcal{F}[\underline{w}] & \text{in } B_r, \\ w^t = \underline{w} & \text{on } \partial B_r, \end{cases} \tag{13.2}$$

where h is given by (5.4). Using the implicit function theorem, see [18] Theorem 5.1, we find that the set of t's for which (13.2) is solvable is open.

Once $C^{1,1}$ global a priori estimates were established in $\overline{B_r}$ then one can deduce that the set of such t's is also closed. Recall that if $\partial\Omega \in C^3, u \in C^4(\Omega) \cap C^3(\overline{\Omega})$ and $\underline{u} \in C^4$ then from global $C^{1,1}$ estimates and the elliptic regularity theory we obtain that $w \in C^{2,\alpha}(\overline{\Omega})$. Therefore the existence of smooth solutions $u^{\pm}_{s,\delta}$ will follow once we establish the global $C^{1,1}$ estimate for w. The latter follows from [6] and Theorem B.

Summarizing, we have that $u^{\pm}_{s,\delta}$ remain locally uniformly smooth in B_r. Letting $s \to 0$ and applying the comparison principle (see Proposition 12.1) we have that $u^-_{0,\delta} \le u \le u^+_{0,\delta}$ and $u^{\pm}_{0,\delta} = u$ on ∂B_r. It follows from the a priori estimates in [14] that $u^{\pm}_{0,\delta}$ are locally uniformly C^2 in B_r for any small $\delta > 0$. After sending $\delta \to 0$ we will conclude the proof of Theorem D.

References

[1] A.D. Aleksandrov, Die innere Geometrie der konvexen Flächen, Berlin, Akademie-Verlag, 1955

[2] R.B. Ash, Measure, Integration and Functional Analysis, Academic Press, New York, 1972

[3] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions. Comm. Pure Appl. Math. 44 (1991), no. 4, 375–417.

[4] L.A. Caffarelli, The Regularity of Mappings with a Convex Potential, Jour. Amer. Math. Soc. Vol. 5, No. 1, (1992), 99–104

[5] C. Gutierrez, Q.Huang, The refractor problem in reshaping light beams. Arch. Ration. Mech. Anal. 193 (2009), no. 2, 423–443.

[6] F. Jiang, N.S. Trudinger, X.-P. Yang, On the Dirichlet problem for Monge-Ampère type equations, to appear in Calc. Var. and PDE

[7] A.L. Karakhanyan, On the regularity of weak solutions to refractor problem, Arm. J. of Math. 2, Issue 1, (2009), 28–37,

[8] A.L. Karakhanyan, Existence and regularity of the reflector surfaces in $\mathbb{R}^{n+1}$, submitted

[9] A.L. Karakhanyan, X.-J. Wang, On the reflector shape design, Journal of Diff. Geometry, 84 (2010), no. 3, 561–610

[10] A.L. Karakhanyan, X.-J. Wang, The reflector design problem, Proceedings of Inter. Congress of Chinese Mathematicians, vol 2, 1-4, 1–24, 2007

[11] M. Kline, Mathematical Thought from Ancient to Modern Times, Volume 1, Oxford University Press, 1972

[12] J. Liu, Light reflection is nonlinear optimization, Calc. Var. Partial Differential Equations 46 (2013), no. 3-4, 861–878.

[13] G. Loeper, On the regularity of solutions of optimal transportation problems, Acta Mathematica, June 2009, Volume 202, Issue 2, pp 241–283

[14] X.N. Ma, N. Trudinger, X.-J. Wang, regularity of potential functions of the optimap transportation problem, Arch. Rat. Mech. Anal. 177 (2005), 151–183

[15] D. Mountford, Refraction properties of Conics. Math. Gazette, 68, 444, (1984), 134-137

[16] A.V. Pogorelov, Monge-Ampère equations of elliptic type, Noordhoff 1964

[17] A.V. Pogorelov, The Minkowski multidimensional problem, John Wiley & Sons Inc 1978

[18] N.S. Trudinger, Lectures on nonlinear elliptic equations of second order, in: Lectures in Mathematical Sciences, The University of Tokyo, 1995

[19] N.S. Trudinger, X.-J. Wang On Strict Convexity and Continuous Differentiability of Potential Functions in Optimal Transportation, Arch. Rational Mech. Anal. 192 (2009) 403–418

[20] J. Urbas, Mass transfer problems, Lecture Notes, Univ. of Bonn, 1998

[21] X.-J. Wang, On the design of a reflector antenna, Inverse Problems, 12, 1996 351–375

Maxwell Institute for Mathematical Sciences and School of Mathematics, University of Edinburgh, King's Buildings, Mayfield Road, EH9 3JZ, Edinburgh, Scotland

E-mail address: aram.karakhanyan@ed.ac.uk