

Absorbing boundary conditions for diffusion equations

L. Halpern^{1,2}, J. Rauch^{3,4,*}

¹ Institut Galilée, Université Paris XIII, F-93430 Villetaneuse, France

² Centre de Mathématiques Appliquées, Ecole Polytechnique, F-91128 Palaiseau, France

³ Department of Mathematics, University of Michigan, Ann Arbor, MI 48104, USA

⁴ Centre de Mathématiques Appliquées and Centre de Mathématiques, Ecole Polytechnique, F-91128, Palaiseau, France

Received December 12, 1992 / Revised version received July 4, 1994

Summary. We construct and analyse a family of absorbing boundary conditions for diffusion equations with variable coefficients, curved artificial boundary, and arbitrary convection. It relies on the geometric identification of the Dirichlet to Neumann map and rational interpolation of $z^{1/2}$ in the complex plane. The boundary conditions are stable, accurate, and practical for computations.

Résumé. Nous introduisons une famille de conditions aux limites absorbantes pour des équations paraboliques à coefficients variables et une frontière quelconque. Elle repose sur l'identification géométrique de l'application Dirichlet à Neumann, et une approximation rationnelle de $z^{1/2}$ dans le plan complexe. Les conditions aux limites obtenues sont stables, précises, et faciles à mettre en œuvre.

Mathematics Subject Classification (1991): 65 P05

1. Introduction

The construction of reliable artificial boundary conditions for the numerical computation of problems in unbounded domains has received much attention (see [4] and references therein). The goal is to impose boundary conditions on an artificial boundary which is close to the domain where one hopes to calculate the solution, and so that the solution of the problem in the reduced domain is a good approximation to the solution of the original problem. The conditions must

Correspondence to: L. Halpern at the first address

* Partially supported by the ONR, NSF, and NATO under grants No. 014 92 J 1245, DMS 90 03256, CRG 890904 respectively

lead to well posed problems which are practical in the sense that they can be implemented numerically. The artificial boundary conditions tend to be nonlocal and the requirement of practicality demands that they be causal, or even better local, in time.

For hyperbolic problems a key idea is to minimize the reflection of waves at the artificial boundary. Engquist and Majda [3] introduced a method which is now standard for the wave equation in a disk. They computed the symbol of the Dirichlet to Neumann map (called simply the *Neumann operator* in the sequel) at the artificial boundary and then found stable, accurate, and practical approximations.

We perform a similar analysis for parabolic equations. For problems in a half space with constant coefficients, problems which are essentially one dimensional, or problems with very small diffusion constants there are related works for example [5–8, 10, 13]. Our approach does not require such special hypotheses in particular it applies to variable coefficients, curved artificial boundaries, and arbitrary advection in the domain.

The analysis consists of four distinct parts. First there is the study of the Neumann operator. Consider the operator

$$Lu \equiv \partial_t u - \sum_{i,j=1}^n a^{ij}(t,x) \partial_i \partial_j u + \text{l.o.t.}$$

Suppose that Ω is a bounded domain in \mathbb{R}^d and S the boundary of the exterior of Ω . If $Lu = 0$ in the exterior of Ω then $\partial_\nu u$ on S is determined by u on S , that is

$$(1.1) \quad \partial u / \partial \nu = -Nu .$$

This defines $N_{\tilde{\Omega}}$ the Neumann operator associated to the equation $Lu = 0$ in the exterior $\tilde{\Omega}$. The minus sign is present because ν is the outward normal to Ω and not $\tilde{\Omega}$. The condition (1.1) is the exactly transparent boundary condition. The Neumann operator is pseudodifferential in $\mathbb{R}_t \times S$. The symbol has an expansion in terms homogeneous in $(\tau^{1/2}, \xi)$

$$(1.2) \quad N(t, x, \tau, \xi) \sim \sum_0^\infty N_{1-j}(t, x, \tau, \xi) .$$

The principal term is easy to describe. The symbol $a^{ij} \xi_i \xi_j$ is for each fixed t a metric on the cotangent bundle $T^*(\mathbb{R}^d)$. This in turn induces a metric on $T^*(S)$. Denote by $|\cdot|$ the associated length. Then

$$(1.3) \quad N_1 = (i\tau + |\xi|^2)^{1/2}$$

the square root taken with positive real part. Particular attention is paid to identifying geometrically the next term N_0 .

The Dirichlet to Neumann operator is nonlocal in space and time. The second step, following Lindmann [12], is to show that using rational approximations of the function $z^{1/2}$ of the form

$$(1.4) \quad z^{1/2} \approx \alpha z + \beta + \sum_{k=1}^m \alpha_k z / (z + d_k) \quad \text{with} \quad \alpha \geq 0, \beta \geq 0, \alpha_k > 0, d_k > 0$$

one can construct stable and computationally feasible approximations to the Neumann operator. Rational approximations as in (1.4) we call *admissible*. For each such approximation we construct four distinct absorbing boundary conditions. Two in Sect. 5 correspond to truncating (1.1) at N_1 and two more in Sect. 6 come from truncation at N_0 .

The two families correspond to distinct ways of using the rational approximation i.e.

$$(1.5) \quad (\partial_t - \Delta_S)^{1/2} \approx \alpha(\partial_t - \Delta_S) + \beta + \sum_{k \geq 1} \alpha_k (\partial_t - \Delta_S)(\partial_t - \Delta_S + d_k)^{-1}$$

and

$$(1.6) \quad \left[(\partial_t - \Delta_S)^{1/2} \approx (-\Delta_S)^{-1/2} \right. \\ \left. \times \left[\alpha(\partial_t - \Delta_S) - \Delta_S \left(\beta + \sum_k \alpha_k (\partial_t - \Delta_S)(\partial_t - (1 + d_k)\Delta_S)^{-1} \right) \right] \right].$$

The second, (1.6), corresponds to the strategy adopted in [10]. A priori it is not clear which of these algorithms is preferable, though in all cases we prefer those which truncate after N_0 as the added precision has small computational cost.

Third, and this is a key step, we show in Sect. 7 that interpolation of the function $z^{1/2}$ at a family of points in $\mathbb{C} \setminus]-\infty, 0]$ which is symmetric with respect to the real axis, by rational functions of exact degree (n, n) or $(n + 1, n)$ (depending on whether the number of interpolation points is odd or even) yields admissible approximations. Note that it is not obvious that such interpolations exist and in fact without the symmetry hypothesis there are counterexamples. That the interpolates have form (1.4) is even less clear.

An interesting question is whether other methods of generating rational approximations, for example least squares might be admissible and possibly better.

Finally in Sect. 8 we present a result which shows that approximating the Dirichlet to Neumann map by its principal symbol or its symbol up to the next order does indeed yield an approximate solution in the sense that the error is smoother than the solution itself.

We would like to call attention to two weaknesses of the analysis. First we have no convergence theorems. We do not provide a sequence of artificial conditions whose precision converges to 100%. On the other hand, exactly this strategy with the same effect has proved to be quite efficient for hyperbolic problems. Second, for hyperbolic problems, approximating the Neumann operator corresponds to avoiding the reflection of singularities at the artificial boundary.

However, for parabolic problems singularities do not propagate to the artificial boundary. For that reason accurately treating singularities, is not as natural. In any event the utility of the methods introduced must be assessed by experience. In addition, there do not appear to be many other options at present.

A detailed numerical study of our conditions in the case of a circular boundary in \mathbb{R}^2 has been carried out by Dubach [2]. We present here a few of his results. He computes solutions of the advection-diffusion equation

$$(1.7) \quad \partial_t u + \partial u / \partial x_1 - \nu \Delta u = 0 \quad \text{in } [0, \infty[\times \mathbb{R}^2$$

$$(1.8) \quad u(0, x) = e^{-\gamma|x|^2}$$

where ν is constant and γ is chosen so that $u(0, x) \geq 10^{-10}$ when $|x| = 2$. We are interested in the values of u in the disk $D(2)$ where $D(r) \equiv \{|x| < R\}$ and we take for computational domains the slightly larger disk $D(2.1)$ and also the disk $D(3)$. Our condition B_{III} corresponding to the choice (1.5) and keeping the terms $j = 0, 1$ from (1.2) leads to a system of equations for u and auxiliary functions $\varphi_k(\theta)$, $k = 0, 1, \dots, m$ on $S \equiv \partial D$. Here $2m + 1$ is the number of interpolation points for $z^{1/2}$. The functions φ_k and most importantly u are determined from (1.7), (1.8) together with

$$(1.9) \quad \nu^{-1} \partial_t \varphi_k - R^{-2} \partial_\theta^2 \varphi + d_k \varphi_k = u, \quad \varphi_k(0, \theta) = 0, \quad 0 \leq k \leq m, \quad d_0 \equiv 0,$$

$$(1.10) \quad -\partial u / \partial r = \left[1/2R + \beta + \sum_1^m \alpha_k - \cos(\theta)/2\nu \right] u - \sum_1^m \alpha_k d_k \varphi_k \\ + (1/2R^3) \partial_\theta^2 \varphi_0 - (\sin(\theta)/2\nu R) \left[\beta \partial_\theta \varphi_0 + \sum_1^m \alpha_k \partial_\theta \varphi_k \right] \quad \text{when } r = R.$$

The somewhat frightening line (1.10) is the absorbing condition for the disk. Notice that computing the φ_k involves solving parabolic equations on $\mathbb{R}_t \times S$ thus with one fewer space variable than for u . In this sense the added computational cost is small.

We take Dubach's example with $m = 3$ and the seven interpolation points

$$0.01692, \quad \pm i0.09254, \quad \pm i0.39899, \quad \pm i0.84195.$$

Note that these numbers are all not far from the origin. This results in the somewhat paradoxical fact that we are approximating high frequency asymptotics at low frequency.

Figure 1 presents the errors and relative errors for $\nu = 2$ and also a comparison with the errors committed when the Dirichlet conditions are imposed on ∂D . The Neumann conditions are much worse than either of these. The error for our

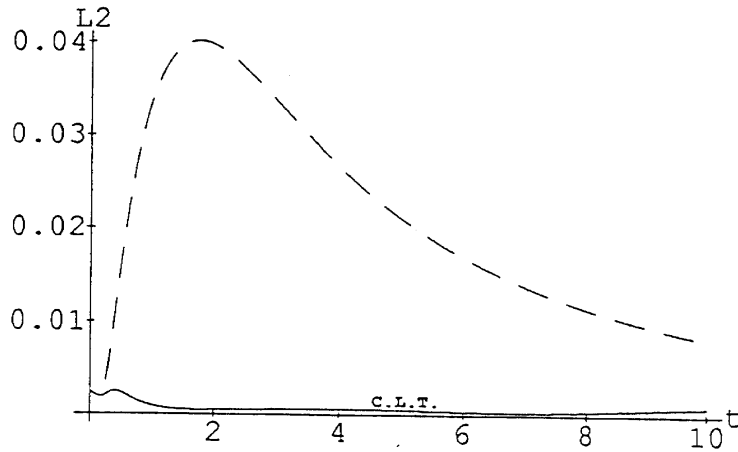


Fig. 1. L^2 error for Dirichlet (dashed) and absorbing (solid) boundary conditions as a function of time

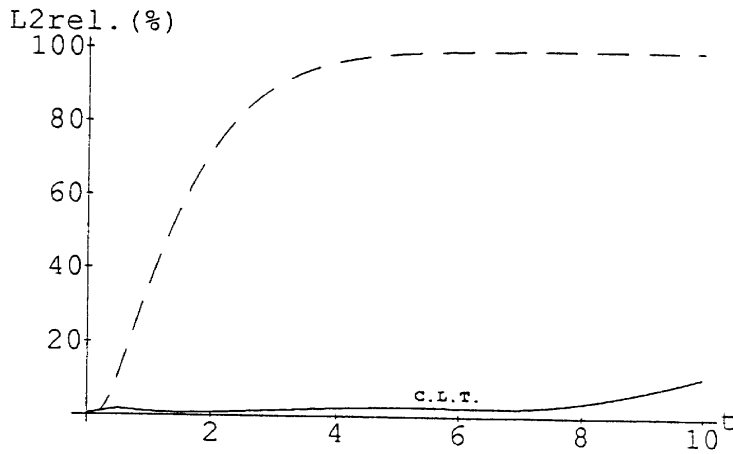


Fig. 2. Relative L^2 error for Dirichlet (dashed) and absorbing (solid) boundary conditions as a function of time

method is much smaller than for Dirichlet yielding relative errors of the order .025 for times up to $t = 5$.

Figure 2 presents a closer look at the relative error for the absorbing condition.

Figure 3 presents the $L^\infty([0, 10] : L^2(|x| \leq 2))$ error as a function of ν showing as expected that the method is better when the diffusion coefficient ν is large.

Figure 4 shows that increasing the size of the disk to $R = 3$ pays very little dividend in this case, supporting our belief that the absorbing conditions are performing as desired. Frankly speaking they perform better than we thought they would!

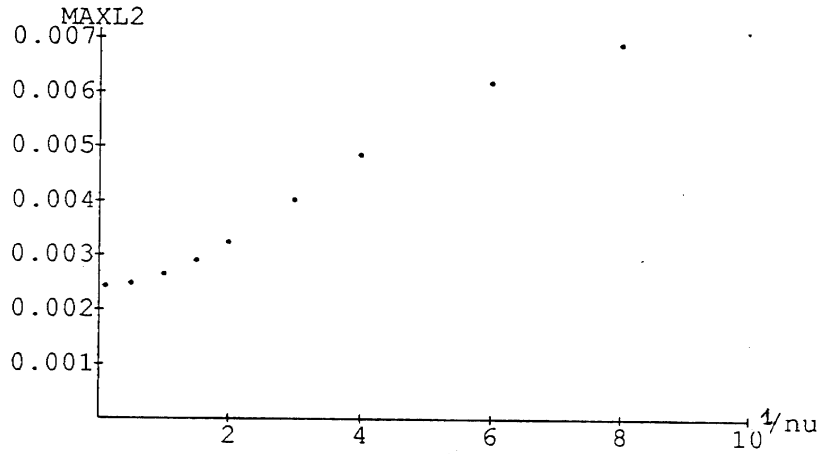


Fig. 3. Dependence of the $L^\infty(L^2)$ error on ν , with $R = 2.1$

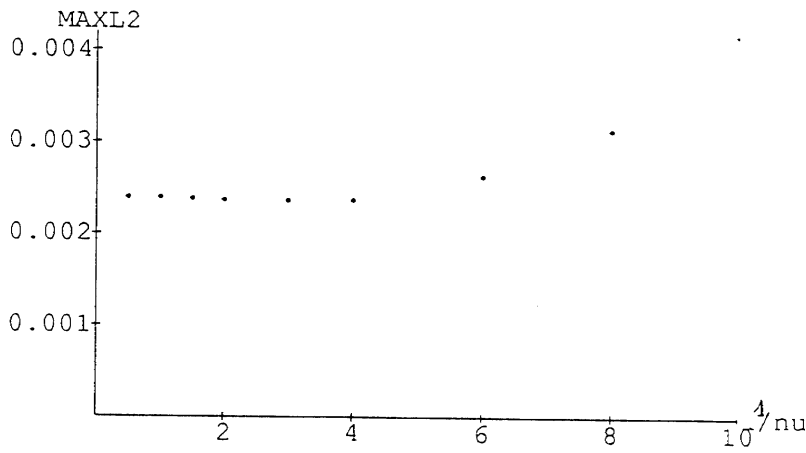


Fig. 4. Dependence of the $L^\infty(L^2)$ error on ν , with $R = 3$

2. A geometric normal form for L

Consider the parabolic operator

$$(2.1) \quad Lu \equiv \partial_t u - \sum_{i,j=1}^n a^{ij}(t,x) \partial_{x_i} \partial_{x_j} u + \text{lower order terms.}$$

The spatial part of the operator is assumed to be smooth and elliptic in the sense that a^{ij} is smooth on \mathbb{R}^{n+1} , $a^{ij} = a^{ji}$, $D^\alpha a^{ij} \in L^\infty(\mathbb{R}^{n+1})$ for all α , and, there is a $\mu > 0$ so that for all $t, x, \xi \in \mathbb{R}^{n+1} \times \mathbb{R}^n$

$$(2.2) \quad \sum a^{ij}(t,x) \xi_i \xi_j \geq \mu(\xi_1^2 + \xi_2^2 + \dots + \xi_n^2).$$

For any x in \mathbb{R}^n , we denote by $T_x(\mathbb{R}^n)$ the linear space of tangent vectors to \mathbb{R}^n at x . The dual space of covectors is denoted $T_x^*(\mathbb{R}^n)$. The corresponding tangent and cotangent bundles are $T(\mathbb{R}^n)$ and $T^*(\mathbb{R}^n)$.

For each t, x , the principal symbol of the spatial part, $\sum a^{ij}(t, x)\xi_i\xi_j$, defines a scalar product on $T_x^*(\mathbb{R}^n)$. The corresponding length is denoted $|\xi|^2 \equiv \langle \xi, \xi \rangle = \sum a^{ij}(t, x)\xi_i\xi_j$.

The associated quadratic form $\langle \xi, \eta \rangle = \sum a^{ij}\xi_i\eta_j$ induces an isomorphism between T_x and T_x^* as follows. Each vector $v \in T_x$ defines a linear functional $\xi \mapsto v(\xi)$ on T_x^* . There is a unique $\eta = \eta(v) \in T_x^*$ such that $v(\xi) = \langle \xi, \eta \rangle$ for all ξ . If $v = v^i\partial/\partial x_i$ (summation convention) the relation defining $\eta(v)$ is $\sum a^{ij}\xi_i\eta_j = v^i\xi_i$. Thus $\eta_i(v) = (a^{-1})_{ij}v^j$.

The isomorphism $v \mapsto \eta(v)$ induces a Riemannian metric $g_{ij}dx^i dx^j$ on the tangent bundle $T(\mathbb{R}^n)$ by

$$g_{ij}v^i v^j \equiv \langle v, v \rangle \equiv \langle \eta, \eta \rangle = a^{ij}\eta_i\eta_j .$$

It follows that

$$(2.3) \quad g(t, x) = a(t, x)^{-1} , \quad \text{that is} \quad g_{ij} = (a(t, x)^{-1})_{ij} .$$

The correspondence $v \mapsto \eta(v)$ is the classical lowering of indices defined by the metric g , and the inverse is raising indices. These operations preserve lengths.

The volume element of the Riemannian metric is equal to

$$d\text{vol} = (\det g_{ij})^{1/2} dx = (\det a)^{-1/2} dx .$$

For a smooth function f , the one form df corresponds to the vector field $\text{grad} f$ by raising indices. One then has the Dirichlet integral

$$D(f) \equiv \int |\text{grad} f|^2 d\text{vol}/2 = \int |df|^2 d\text{vol}/2 = \int a^{ij}\partial_i f \partial_j f d\text{vol}/2 .$$

The Laplace-Beltrami operator for a Riemannian metric, g , is defined by

$$(2.4) \quad \text{Variation of } D(f) \text{ in direction } \varphi \equiv - \int \varphi \Delta_g f d\text{vol} .$$

The variation on the left is the Frechet derivative

$$\begin{aligned} (d/d\varepsilon)D(f + \varepsilon\varphi)|_{\varepsilon=0} &= \int \langle \text{grad} \varphi, \text{grad} f \rangle d\text{vol} \\ &= \int a^{ij}\partial_i f \partial_j \varphi d\text{vol} . \end{aligned}$$

The reader is warned that this is the analyst's sign convention, the opposite sign being the usual choice in geometry. The analyst's choice yields

$$(2.5) \quad \Delta_g v = \sum_{i,j=1}^n (\det g)^{-1/2} \partial_{x_i} ((\det g)^{1/2} a^{ij} \partial_{x_j} v) .$$

The principal part of this operator is equal to $\sum a^{ij} \partial_i \partial_j$ which is equal to the principal part of the spatial part of L . In the sequel g^{ij} will denote $(g^{-1})_{ij} = a^{ij}$. This underlines the connection between L and the geometry defined by the metric g induced by a .

So far we have observed that

$$(2.6) \quad L = \partial_t - \Delta_g + \text{lower order terms.}$$

Under coordinate changes $\mathbf{t} = t$, $\mathbf{x} = \mathbf{x}(t, x)$ which preserve the time variable, Δ_g has invariant meaning as does the degree zero term in L which has coefficient given by $s(t, x) \equiv L1$.

Then $L + \Delta_g - s$ is a first order differential operator which annihilates constants. Thus it is a vector field in t, x . The coefficient of $\partial/\partial t$ is equal to one. The latter condition is preserved by the above changes of variables.

Proposition 2.1. *If $g = a^{-1}$ and $s \equiv L1$, then there is a unique time dependent vector field V on $\mathbb{R}_{t,x}^{1+n}$*

$$(2.7) \quad V = \partial/\partial t + \sum v^j(t, x) \partial/\partial_j ,$$

such that

$$(2.8) \quad L = V - \Delta_g + s .$$

Note that neither ∂_t nor $\sum v^j \partial_j$ are well defined operators under the above coordinate changes. For example

$$\partial/\partial t = \partial/\partial \mathbf{t} + \sum (\partial \mathbf{x}_j / \partial t) \partial/\partial \mathbf{x}_j \neq \partial/\partial \mathbf{t}$$

except if the change of variables is independent of time.

3. The symbol of the Neumann operator

Suppose that $\Omega \subset \mathbb{R}^n$ is an open set which together with its boundary is an embedded submanifold with boundary. Let S denote the boundary of Ω .

Since $\mathbb{R} \times S$ is noncharacteristic for L it follows that if $t, x \in \mathbb{R} \times S$ and u is a square integrable solution of $Lu = 0$ on $B_r(t, x) \cap (\mathbb{R} \times \Omega)$, then all the derivatives of u have well defined traces on $B_r(t, x) \cap (\mathbb{R} \times S)$ ($B_r(t, x)$ is the ball of radius r centered at point (t, x)).

For $t, x \in \mathbb{R} \times S$, let $\nu(t, x)$ be the unit outward pointing normal to Ω for the metric $g_{ij}(t, x)$. The normal derivative $\partial_\nu u$ is then defined to be $du(\nu) = \langle \text{grad } u, \nu \rangle_g$. The previous paragraph shows that $\partial_\nu u|_S$ and $u|_S$ are well defined distributions on a neighborhood of t, x in S . Standard regularity theorems for the Dirichlet boundary condition on $\mathbb{R} \times \bar{\Omega}$ for the operator L assert that if Lu is

smooth on $B_r(t, x) \cap (\mathbb{R} \times \bar{\Omega})$ and $u|_S$ is smooth on a neighborhood of t, x , then u is smooth on an $\mathbb{R} \times \bar{\Omega}$ neighborhood of t, x .

Consider the Dirichlet problem for the operator L . For $h \in C_0^\infty(\mathbb{R} \times S)$, let u be the unique solution of

$$(3.1) \quad Lu = 0 \quad \text{in} \quad \mathbb{R} \times \Omega, \quad u|_{\mathbb{R} \times S} = h, \quad u \equiv 0 \quad \text{for} \quad t \ll 0.$$

The map $N : h \mapsto \partial_\nu u|_S$ is called the *Neumann operator* or Dirichlet to Neumann operator for the domain Ω . For emphasis we sometimes write N_Ω .

If $h \equiv 0$ for $t < t_0$ then u and therefore Nh vanish for $t < t_0$. Thus N is *causal* in the sense that *the values of Nh in $t < T$ are determined by the values of h in $t < T$.*

The regularity theorem asserts that the operator N is pseudolocal in the sense that

$$(3.2) \quad \text{singsupp}(Nu) \subset \text{singsupp}(u).$$

Even more is true. The operator N is a pseudodifferential operator with symbol

$$(3.3) \quad N(t, x, \tau, \xi) \sim \sum_0^\infty N_{1-j}(t, x, \tau, \xi)$$

where N_k is *homogeneous of degree k in $\tau^{1/2}, \xi$* in the sense that

$$(3.4) \quad N_k(t, x, \lambda^2 \tau, \lambda \xi) = \lambda^k N_k(t, x, \tau, \xi) \quad (\forall (\tau, \xi) \neq 0, \lambda > 0).$$

N_k belongs to Hormander's symbol classes $S_{1/2,0}^k$ (see [16]). A more refined symbol class adapted to heat equations is described at the end of Sect. 4.

To compute the symbol N , the key idea is to use geodesic normal coordinates. For $p \in S$, let $\gamma(s, p)$ be the unique geodesic for the metric g such that $\gamma(0, p) = p$ and $\gamma'(0, p) = \nu(p)$. Let $x' \equiv (x_1, \dots, x_{n-1})$ denote local coordinates for a neighborhood of p in S . The map $(x', x_n) \mapsto \gamma(x_n, p(x'))$ defines a local coordinate system for a neighborhood of p in \mathbb{R}^n . Since ν is an outward normal, the interior of Ω corresponds to $x_n < 0$. The metric g has the form

$$(3.5) \quad g = \sum_{\alpha, \beta=1}^{n-1} g_{\alpha\beta} dx^\alpha dx^\beta + (dx_n)^2.$$

Here we introduce the convention that *Greek indices run from 1 to $n-1$ and Roman indices run from 1 to n .*

Since $g_{nj} = 0$ for $j < n$ and $g_{nn} = 1$ the Laplace-Beltrami operator takes the form

$$(3.6) \quad \Delta_g u \equiv \rho^{-1} \partial_{x_n} (\rho \partial_{x_n} u) + \sum_{\alpha\beta} \rho^{-1} \partial_{x_\alpha} (\rho g^{\alpha\beta} \partial_{x_\beta} u)$$

where

$$\rho \equiv (\det g_{ij})^{1/2} = (\det g_{\alpha\beta})^{1/2}$$

so the volume element is $\rho(x)dx$ and $g^{\alpha\beta} = (g^{-1})_{\alpha\beta}$.

With the standard notation $D \equiv -i\partial \equiv (D_1, \dots, D_n) \equiv (D', D_n)$, the operator L takes the form

$$(3.7) \quad L = D_n^2 + i(-\rho^{-1}\partial_n\rho + v_n)D_n + Q(t, x, D') + iD_t,$$

$$Q \equiv Q_2 + Q_1 + Q_0,$$

$$(3.8) \quad Q_2 \equiv \sum g^{\alpha\beta} D_\alpha D_\beta,$$

$$Q_1 \equiv \sum_\beta \left\{ v_\beta - \sum_\alpha \rho^{-1} \partial_\alpha (\rho g^{\alpha\beta}) \right\} iD_\beta, \quad Q_0 = s.$$

The calculation of N proceeds by factoring L , a procedure which is now standard (see [3, 11, 15, 16]).

Theorem 3.1. *Suppose that geodesic normal coordinates are introduced as above. Then there are tangential pseudodifferential operators $A(t, x, D_t, D')$, $B(t, x, D_t, D')$ in*

$$C^\infty \left([1 - \delta, \delta]_{x_n} : OpS_{1/2,0}^1(\mathbb{R}_{t,x'}^n \times \mathbb{R}_{\tau,\xi'}^n) \right)$$

with symbols

$$(3.9) \quad A(t, x, \tau, \xi') \sim \sum_{j \geq 0} A_{1-j}(t, x, \tau, \xi'),$$

$$B(t, x, \tau, \xi') \sim \sum_{j \geq 0} B_{1-j}(t, x, \tau, \xi')$$

with A_k, B_k homogeneous of degree k in $\tau^{1/2}, \xi$ satisfying

$$(3.10) \quad L = (D_n + A)(D_n + B) \text{ mod } C^\infty([1 - \delta, \delta] : OpS_{1/2,0}^{-\infty}).$$

$$(3.11) \quad \text{Im } B_1 > 0.$$

The homogeneous terms A_k and B_k are uniquely determined. For $k = 1, 0$ they are given by formulas (3.12), (3.13), (3.15).

Proof. To show the uniqueness expand the right hand side of (3.10)

$$(D_n + A)(D_n + B) = D_n^2 + (A + B)D_n + C(t, x, D_{t,x'})$$

$$C = AB + [D_n, B] \sim -i\partial B / \partial x_n + \sum (\partial_{\tau, \xi'}^\alpha A)(D_{t,x'}^\alpha B) / \alpha!.$$

The A_k and B_k are determined by the conditions

$$A + B = i(-\rho^{-1}\partial_n\rho + v_n), \quad \text{and}$$

$$C = Q + iD_t \text{ mod } C^\infty([1 - \delta, \delta] : OpS_{1/2,0}^{-\infty}).$$

The first condition is satisfied if and only if

$$(3.12) \quad A_k + B_k = 0 \text{ if } k \neq 0 \text{ and } A_0 + B_0 = i(-\rho^{-1}\partial_n\rho + v_n) .$$

Computing the leading terms of $C \sim C_2 + C_1 + \dots$ yields

$$\begin{aligned} C_2 &= A_1 B_1 \\ C_1 &= A_0 B_1 + A_1 B_0 + D_n B_1 + \sum (\partial_{\xi_\alpha} A_1)(D_{x_\alpha} B_1) . \end{aligned}$$

In making the second calculation recall that the τ derivative of a function homogeneous of degree k in $\tau^{1/2}$, ξ' is homogeneous of degree $k - 2$ so the $(\partial_\tau A_1)(D_t B_1)$ term is of order zero and does not contribute to C_1 .

The degree 2 term in $C = Q + iD_t$ combined with (3.12) for $k = 1$ yields

$$(3.13) \quad -(B_1)^2 = i\tau + \sum g^{\alpha\beta} \xi_\alpha \xi_\beta \quad \text{so} \quad B_1 = i(i\tau + |\xi'|^2)^{1/2}$$

the square root being the one with positive real part so that B_1 satisfies (3.11).

Identities (3.12) for $k = 0, 1$ show that the degree one term in C is given in terms of B_0 and B_1 by

$$(3.14) \quad \begin{aligned} C_1 &= (-2B_1)B_0 + i(-\rho^{-1}\partial_n\rho + v_n)B_1 + D_n B_1 \\ &\quad - \sum (\partial_{\xi_\alpha} B_1)(D_{x_\alpha} B_1) . \end{aligned}$$

Setting this equal to Q_1 yields

$$(3.15) \quad B_0 = i(-\rho^{-1}\partial_n\rho + v_n)/2 + \frac{-Q_1 + D_n B_1 - \sum (\partial_{\xi_\alpha} B_1)(D_{x_\alpha} B_1)}{2B_1}$$

where division by B_1 is justified since B_1 is nowhere zero in $\tau, \xi' \neq 0$.

Continuing in this fashion uniquely determines A_k and B_k .

Choose $A, B \in C^\infty([-\delta, \delta]; S_{1/2,0}^m)$ satisfying (3.9). The calculus of pseudodifferential operators then implies that

$$\sigma(L - (D_n + A)(D_n + B)) \in C^\infty([-\infty, 0] : S_{1/2,0}^{-\infty})$$

which proves (3.10). \square

Remark. Neither B_0 nor B_1 depends on the zeroth order term $s(t, x)$.

Theorem 3.2 *The Neumann operator N_Ω is equal to $-iB$ modulo a smoothing operator.*

Proof. Thanks to (3.11) the final value problem

$$(3.16) \quad D_n U + BU = 0 \quad \text{in } x_n \leq 0, \quad U|_{x_n=0} = h$$

has for any $h \in H^s(\mathbb{R}_{t,x'}^n)$ a unique solution $U \in C([-\delta, 0]_{x_n} : H^s(\mathbb{R}_{t,x'}^n))$.

Thanks to (3.10), this solution U satisfies $LU \in C^\infty(-\delta < x_n \leq 0)$. Thus $L(U - u) \in C^\infty$ and $U - u$ vanishes at $x_n = 0$. The Local Regularity Theorem for the Dirichlet problem for L implies that $U - u \in C^\infty(-\delta < x_n \leq 0)$.

Since the metric has the form (3.5) the conormal derivative at the boundary is $\partial/\partial x_n$. Therefore at $x_n = 0$,

$$Nh \equiv \partial u / \partial x_n = \partial U / \partial x_n + C^\infty = -iBU + C^\infty = -iBh + C^\infty . \quad \square$$

An interesting special case occurs when $s(t, x)$ and the $v_j(t, x)$ vanish and the operator L in geodesic normal coordinates has coefficients independent of time. This is the heat equation $\partial_t - \Delta$ where Δ is the Laplace-Beltrami operator on the product Riemannian manifold $S \times \mathbb{R}_+$ with metric equal to $g_S + g_{\mathbb{R}} = g_S + dx_n^2$. The heat operator has the form

$$(3.17) \quad L = \partial_t - \Delta_S - \partial_n^2 .$$

The Neumann operator in this case is denoted N_0 .

Proposition 3.3. *For the heat equation on the product $\mathbb{R}_t \times (S \times \mathbb{R}_+)$ the Neumann operator N_0 satisfies $N_0^2 = \partial_t - \Delta_S$.*

Proof. If u is a solution of $Lu = 0$ then $w \equiv \partial u / \partial x_n$ satisfies $Lw = 0$. Therefore at $x_n = 0$,

$$\partial_n^2 u = \partial_n w = N_0(w|_{\mathbb{R} \times S}) = N_0(\partial_n u|_{\mathbb{R} \times S}) = N_0(N_0(u|_{\mathbb{R} \times S})) .$$

On the other hand, $\partial_n^2 u = (\partial_t - \Delta_S)u$ and the result is proved. \square

Proposition 3.3 motivates the following definition.

Definition. *The operator $(\partial_t - \Delta_S)^{1/2}$ is defined to be the Neumann operator associated to heat equation on $\mathbb{R}_t \times (S \times \mathbb{R}_+)$. It is denoted N_0 .*

Then,

$$(3.18) \quad \begin{aligned} \sigma((\partial_t - \Delta_S)^{1/2}) &= \sigma(N_0) = -iB_1 - iB_0 + \text{l.o.t.} \\ &= -iB_1 + \frac{\sum_\alpha \rho^{-1} \partial_\alpha (\rho g^{\alpha\beta}) \xi_\beta + i \sum (\partial_{\xi_\alpha} B_1)(D_{x_\alpha} B_1)}{2B_1} + \text{l.o.t.} \end{aligned}$$

where B_1 is given by (3.13).

Remark. The operator $\partial_t - \Delta_S$ on $\mathbb{R}_t \times S$ has many square roots. The one singled out above is characterized by the following three properties

- i.) It commutes with operators that commute with $\partial_t - \Delta_S$.
- ii.) It is pseudodifferential with classical symbol expansion.
- iii.) It has principal symbol with nonnegative real part.

Several of the terms in (3.15) have already been identified. Write $V - \partial/\partial t$ as a sum of its tangential and normal components

$$(3.19) \quad \begin{aligned} V - \partial/\partial t &= V_{\text{tan}} + V_{\text{nor}} , \quad V_{\text{nor}} = \langle (V - \partial/\partial t), \nu \rangle \nu , \\ \sigma(V_{\text{tan}}) &= \sum v_\alpha(t, x) i \xi_\alpha . \end{aligned}$$

In geodesic normal coordinates, $V_{\text{nor}} = v_n \partial_n$. The next result summarizes the information in Theorem 3.2 and equations (3.13), (3.15), (3.18).

Theorem 3.4. *The Neumann operator N_Ω satisfies*

$$(3.20) \quad \begin{aligned} \sigma(N_\Omega) = & \sigma((\partial_t - \Delta_S)^{1/2}) + i(\sigma(V_{\text{tan}}) - D_n B_1)/2B_1 \\ & + (-\rho^{-1} \partial_n \rho + v_n)/2 \bmod S_{1/2,0}^{-1/2}. \end{aligned}$$

Question. The operator $\sigma((\partial_t - \Delta_S)^{1/2})$ is determined by a calculation inside $\mathbb{R}_t \times S$ and does not depend on the embedding $S \hookrightarrow \mathbb{R}^n$. The order 1 part of this operator has a clear interpretation in terms of the first fundamental form of S . We do not know a geometric interpretation of the zeroth order term in the symbol of $\sigma((\partial_t - \Delta_S)^{1/2})$. We are equally ignorant with regard to the zeroth order term in $\sigma(-\Delta_S)^{1/2}$.

4. Geometric identification $B_1 + B_0$

The identity

$$(4.1) \quad B_1 = i(i\tau + g^{\alpha\beta} \xi_\alpha \xi_\beta)^{1/2} = i(i\tau + |\xi'|^2)^{1/2}$$

gives B_1 in terms of the first fundamental form of S . In this sense it is geometric. In this section, we compute a geometric formula for the symbol of B_0 which depends, in addition, on the second fundamental form and V . Put another way, the goal of this section is to identify geometrically the two terms in equation (3.20) which involve derivatives with respect to x_n .

The formula for B_0 involves no time derivatives. It is determined by the geometry of $S \hookrightarrow \mathbb{R}^n$, g and $V(t, \cdot)$ for t fixed and not on the way the geometry depends on time. We begin by recalling some of the basic notions concerning the local Riemannian geometry of the embedding $S \hookrightarrow \mathbb{R}^n$, g (see [9]).

In geodesic normal coordinates for a collar about S , the vectors $\partial/\partial x_j$, $1 \leq j \leq n$ form a basis for the tangent space to \mathbb{R}^n , and $\partial/\partial x_n$ is orthogonal to the others. The first $n-1$ vectors when $x_n = 0$ are a basis for the tangent space to S .

Denote by \bar{D} the Riemannian connection for \mathbb{R}^n , g . Then

$$(4.2) \quad \bar{D}_{\partial_i} \partial_j = \sum_{k=1}^n \bar{\Gamma}_{ji}^k \partial_k.$$

The time dependent Christoffel symbols $\bar{\Gamma}$ are given by

$$(4.3) \quad \bar{\Gamma}_{ji}^k = \frac{1}{2} \sum_{m=1}^n g^{km} (\partial g_{im}/\partial x_j + \partial g_{jm}/\partial x_i - \partial g_{ji}/\partial x_m).$$

The metric g on \mathbb{R}^n induces a metric on S which in turn determines a Riemannian connection, denoted D , on S

$$(4.4) \quad D_{\partial_\alpha} \partial_\beta = \sum_{\gamma=1}^{n-1} \Gamma_{\beta\alpha}^\gamma \partial_\gamma$$

where the Christoffel symbols $\Gamma_{\beta\alpha}^\gamma$ are given by

$$(4.5) \quad \Gamma_{\beta\alpha}^\gamma = \frac{1}{2} \sum_{\mu=1}^{n-1} g^{\gamma\mu} (\partial g_{\alpha\mu} / \partial x_\beta + \partial g_{\beta\mu} / \partial x_\alpha - \partial g_{\beta\alpha} / \partial x_\mu).$$

In the geodesic normal coordinates,

$$g = \begin{bmatrix} g_{\alpha\beta} & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad g^{-1} = \begin{bmatrix} (g_{\alpha\beta})^{-1} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} g^{\alpha\beta} & 0 \\ 0 & 1 \end{bmatrix}.$$

Note that ∂_n is a unit normal to S with respect to $\langle \cdot, \cdot \rangle_g$ so the Weingarten map $W : T(S) \rightarrow T(S)$ is defined by $WX = \bar{D}_X \partial_n$. W is a symmetric linear operator with respect to the scalar product $\langle \cdot, \cdot \rangle_g$ for $T(S)$. The second fundamental form of S is defined as the quadratic form $II(X, Y) \equiv \langle WX, Y \rangle_g$ on $T(S)$.

The formula of Gauss reads

$$(4.6) \quad \bar{D}_X Y = D_X Y - \langle WX, Y \rangle_g \partial_n \quad \forall X, Y \text{ tangent to } S.$$

This gives the orthogonal decomposition of $D_X Y$ into parts tangent to S and orthogonal to S and also relates the connections on S, g and \mathbb{R}^n, g . Taking $X = \partial_\alpha$ and $Y = \partial_\beta$ formula (4.2) also gives an orthogonal decomposition, so the normal parts must be equal

$$(4.7) \quad \bar{\Gamma}_{\alpha\beta}^n = -\langle W \partial_\alpha, \partial_\beta \rangle_g = -II(\partial_\alpha, \partial_\beta).$$

In the sum (4.2) defining the coefficients $\bar{\Gamma}_{ji}^n$, the term g^{nm} vanishes unless $m = n$ in which case it is equal to one. In the $n = m$ summand $g_{im} = g_{jm} \equiv 0$ which yields $\bar{\Gamma}_{\alpha\beta}^n = -(\partial g_{\beta\alpha} / \partial x_n) / 2$. Together with (4.7) this yields

$$(4.8) \quad \partial g_{\alpha\beta} / \partial x_n |_{x_n=0} = 2 \langle W \partial_\alpha, \partial_\beta \rangle_g = 2II(\partial_\alpha, \partial_\beta).$$

Introduce the matrix $W_{\alpha\gamma}$ of W in the basis ∂_α , $W \partial_\alpha \equiv \sum W_{\alpha\gamma} \partial_\gamma$. Then

$$(4.9) \quad \langle W \partial_\alpha, \partial_\beta \rangle_g = \sum_{\gamma} W_{\alpha\gamma} \langle \partial_\gamma, \partial_\beta \rangle_g = \sum_{\gamma} W_{\alpha\gamma} g_{\gamma\beta} = (Wg)_{\alpha\beta},$$

where Wg is the matrix product. Note that the matrix $W_{\alpha\beta}$ need not be symmetric. The symmetry of the Weingarten map is equivalent to the symmetry of the matrix $(Wg)_{\alpha\beta}$. Identities (4.8) and (4.9) yield the matrix equation $\partial g / \partial x_n = 2Wg$.

Now consider the term $\rho^{-1} \partial_n \rho$ in (3.15). Denote by g_1, \dots, g_n the columns of the matrix g_{ij} . We have just shown that $\partial_n g_j = 2Wg_j$. To compute $\rho^{-1} \partial_n \rho$ note that

$$(4.10) \quad \partial_n \rho = \partial_n (\det g)^{1/2} = \frac{1}{2} (\det g)^{-1/2} \partial_n (\det g), \quad \text{and}$$

$$\begin{aligned}
\partial_n(\det g) &= \sum_{j=1}^n \det(g_1, \dots, g_{j-1}, \partial_n g_j, g_{j+1}, \dots, g_n) \\
(4.11) \quad &= 2 \sum_{j=1}^n \det(g_1, \dots, g_{j-1}, W g_j, g_{j+1}, \dots, g_n) .
\end{aligned}$$

The map sending the g_k to the right hand side is an element of $\Lambda^n(\mathbb{R}^n)$. Since $\Lambda^n(\mathbb{R}^n)$ has dimension equal to 1, there is a constant $c(W)$ such that the right hand side is equal to $c(W)\det(g)$. Taking $g = I$, yields $c(W) = 2\text{Tr}(W)$. Thus

$$(4.12) \quad \rho^{-1} \partial_n \rho = (\det g)^{-1/2} \frac{1}{2} (\det g)^{-1/2} (2\text{Tr}(w)) (\det g) = \text{Tr}(W) .$$

The mean curvature H of S is by definition equal to $\text{Tr}(W)/\dim(S)$, so

$$(4.13) \quad \rho^{-1} \partial_n \rho \equiv (n-1)H(t, x) .$$

Examples. The mean curvature is defined by the outward pointing normal to Ω . Thus if Ω is the interior (resp. exterior) of the ball of radius r in \mathbb{R}^n , then $H = 1/r$ (resp. $-1/r$).

The computation for the $D_n B_1$ term in (3.15) is similar. One has

$$(4.14) \quad \partial_n (\text{i}\tau + g^{\alpha\beta} \xi_\alpha \xi_\beta)^{1/2} = \frac{1}{2} (\text{i}\tau + g^{\alpha\beta} \xi_\alpha \xi_\beta)^{-1/2} \partial_n (g^{\alpha\beta} \xi_\alpha \xi_\beta) .$$

The derivative of the matrix $g^{\alpha\beta}$ of g^{-1} is computed as

$$\partial_n g^{-1} = -g^{-1} (\partial_n g) g^{-1} = -g^{-1} 2W g g^{-1} = -2g^{-1} W$$

so

$$(4.15) \quad \partial_n (g^{\alpha\beta} \xi_\alpha \xi_\beta) = -2g^{\alpha\gamma} W_{\gamma\beta} \xi_\alpha \xi_\beta .$$

When $x_n = 0$, formula 4.15 has an elegant geometric interpretation. In the right hand side insert the index lowering isomorphism $\xi_\alpha = g_{\mu\alpha} v^\mu$ from $T(S)$ to $T^*(S)$ to find for $\xi \in T^*(S)$

$$\partial_n (g^{\alpha\beta} \xi_\alpha \xi_\beta) = -2g^{\alpha\gamma} W_{\gamma\beta} g_{\mu\alpha} v^\mu g_{\chi\beta} v^\chi .$$

Since $g^{\alpha\gamma} g_{\mu\alpha} = \delta_\mu^\gamma$ this yields

$$(4.16) \quad \partial_n (g^{\alpha\beta} \xi_\alpha \xi_\beta) = -2g_{\beta\chi} W_{\gamma\beta} v^\gamma v^\chi \equiv -2II(v, v) \equiv -2II(\xi, \xi) ,$$

where the second fundamental form is transported from $T(S)$ to $T^*(S)$ by raising indices.

Equations (4.14) and (4.16) yield

$$(4.17) \quad \text{i}D_n B_1(t, x', \tau, \xi')/B_1 = -II(\xi', \xi')/(\text{i}\tau + |\xi'|^2) .$$

Combining (4.3), (4.13) and (4.17) proves the following formula.

Theorem 4.1 *In local coordinates x' for S the symbol of the Neumann operator N_Ω is given, modulo $S_{1/2,0}^{-1/2}(\mathbb{R}_t \times \mathbb{R}_{x'}; \mathbb{R}_\tau \times \mathbb{R}_{\xi'})$, by*

$$(4.18) \quad \sigma((\partial_t - \Delta_S)^{1/2}) + \frac{-(n-1)H + \langle v - \partial/\partial t, \nu \rangle}{2} + \frac{\sigma(v_{\tan})}{2\Psi^{1/2}} + \frac{II(\xi', \xi')}{2\Psi}$$

where $\Psi \equiv i\tau + \langle \xi', \xi' \rangle_g^2$.

The error estimate implicit in Theorem 4.1 is not optimal. For example, operators with symbols in $S_{1/2,0}^{-1/2}$ map H^s to $H^{s+1/2}$ while the difference between the Neumann operator and the operator with symbol given by (4.18) gains 1/2 a derivative in t and a full derivative in x' . To get this more precise result it is natural to introduce the weight function

$$(4.19) \quad \Phi(\tau, \eta) \equiv (1 + |\tau|^2)^{1/4} + (1 + |\eta|^2)^{1/2}, \quad \eta \equiv \xi'$$

and the inhomogeneous Sobolev space

$$(4.20) \quad H_{\text{Heat}}^s = \{u \in s'(\mathbb{R}^n) : |\hat{u}(\tau, \eta)|\Phi^s \in L^2(\mathbb{R}_{\tau, \eta}^n)\}.$$

Here s is a real number and not the zero order term in L . These spaces are natural for the heat equation.

We use the Beals-Fefferman calculus of pseudodifferential operators with dual weight $\varphi \equiv 1$ [1]. The Beals-Fefferman symbol class $s_{\Phi, \varphi}^{m \ln(\Phi)}$ is defined by the estimates

$$|\partial_{t,x}^\alpha \partial_{\tau, \eta}^\beta p(t, x, \tau, \eta)| \leq c_{\alpha, \beta} \Phi^{m - |\beta|}.$$

For brevity we denote this symbol class as S_{Heat}^m . A function homogeneous of degree m in $\tau^{1/2}, \eta$ for large τ, η defines an element of S_{Heat}^m and the operators in OpS_{Heat}^m map H_{Heat}^s continuously to H_{Heat}^{s-m} .

The difference between the Neumann operator and the operator with symbol given by (4.18) belongs to OpS_{Heat}^{-1} . These remarks suffice to justify the more precise result indicated after the statement of Theorem 4.1. We summarize the basic facts relating the Neumann operator and this calculus.

Theorem 4.2. *The operator $B(t, x, D_t, D_{x'})$ in Theorem 3.1 belongs to $C^\infty(\mathbb{R}^n; OpS_{\text{Heat}}^1)$. For any $m \leq 1$, the difference $B - \sum_{k \geq m} B_k$ with B_k from Theorem 3.1, belongs to $C^\infty(\mathbb{R}^n; OpS_{\text{Heat}}^{m-1})$. In particular, the symbol of N_Ω is equal to (4.18) modulo S_{Heat}^{-1} .*

5. First order absorbing conditions

Consider the numerical computation of a solution u to $Lu = 0$ in all of \mathbb{R}_x^n . Usually one truncates the domain to a practical (hence bounded) domain $\Omega \subset\subset \mathbb{R}^n$. The differential equation is then supplemented with boundary conditions at the artificial boundary $S \equiv \partial\Omega$.

The exact solution u satisfies $Lu = 0$ on all of the exterior $\mathbb{R} \times \tilde{\Omega}$ so

$$(5.1) \quad \partial u / \partial \nu_{\tilde{\Omega}} = N_{\tilde{\Omega}}(u|_{\mathbb{R} \times S}).$$

Since the outward normal to the exterior is opposite the normal to Ω it follows that u in $\mathbb{R} \times \tilde{\Omega}$ satisfies

$$(5.2) \quad \partial u / \partial \nu_{\Omega} = -N_{\tilde{\Omega}}(u|_{\mathbb{R} \times S}).$$

This is called the *transparent boundary condition* in $\mathbb{R} \times \Omega$, since the exact full space solution satisfies this condition.

The strategy is to impose boundary conditions at the artificial boundary which approximate (5.2). That is solve

$$(5.3) \quad Lu_{\text{appr}} = 0 \quad \text{in} \quad \mathbb{R}_+ \times \Omega, \quad \text{and} \quad \partial u_{\text{appr}} / \partial \nu_{\Omega} = Bu_{\text{appr}} \quad \text{on} \quad \mathbb{R}_+ \times S,$$

where $B \approx -N_{\tilde{\Omega}}$ in some sense. Practical considerations suggest that the approximate boundary condition must be *causal* and *stable*. *Stability* means that (5.3) defines a well posed initial value problem. *Causality* means that the values of Bu in $t < T$ depend only on the values of u in $t < T$. This is needed so that a marching scheme can be implemented. Note that we do not insist on locality in x or t . Some standard approximations are nonlocal.

Recall that the principal symbol of $N_{\tilde{\Omega}}$ is equal to $(i\tau + |\xi'|^2)^{1/2}$ where the length square is given by the metric $g^{\alpha\beta} \xi_{\alpha} \xi_{\beta}$ induced by the spatial part of L . In this section two families of algorithms are proposed. Both are generated by taking rational approximations to the function $(i\tau + |\xi'|^2)^{1/2}$. Both methods have the feature that their stability is proved directly by integration by parts as opposed to verifying Lopatinski conditions.

Both families are generated by rational approximations to $z^{1/2}$ of the form

$$(5.4) \quad F(z) \equiv z^{1/2} \approx R(z) = \alpha z + \beta + \sum_{k=1}^m \alpha_k z / (z + d_k),$$

$\alpha \geq 0, \beta \geq 0, \alpha_k > 0, d_k > 0.$

Note that one can equally well write

$$R(z) = \alpha z + b + \sum_{k=1}^m c_k / (z + d_k)$$

but the c_k may be negative even when β and α_k are positive. *The special form (5.4) is crucial in proving stability.*

Section 7 is devoted to showing that approximations of this sort are generated by rational interpolation of the function $F(z)$ by rational functions of exact type (p, q) with $q \leq p \leq q + 1$ provided that the interpolation points lie in $\text{Re}(z) \geq 0$ and are symmetric with respect to the real axis. If the number of interpolation points is odd the approximant of type (q, q) satisfies $\alpha = 0$ and this yields a somewhat simpler form for the algorithm which follows.

The choice of the interpolation points is the subject of research, both theoretical and numerical, of E. Dubach [2].

The approximation (5.4) yields the formal identity

$$(\partial_t - \Delta_S)^{1/2} \approx \alpha(\partial_t - \Delta_S) + \beta + \sum_k \alpha_k (\partial_t - \Delta_S)(\partial_t - \Delta_S + d_k)^{-1}.$$

The operator on the right must be applied to $u|_{\mathbb{H} \times S}$. There is a natural causal realization of $(\partial_t - \Delta_S + d_k)^{-1}$ given by solving auxiliary problems introduced by Lindmann [12]. For $1 \leq k \leq m$ introduce $\varphi_k = (\partial_t - \Delta_S + d_k)^{-1}u$ solution of

$$(5.5) \quad (\partial_t - \Delta_S + d_k)\varphi_k = u \quad \text{for } t \geq 0, \quad \varphi_k|_{t=0} = 0, \quad k = 1, 2, \dots, m.$$

Note that this is a parabolic initial value problem on $[0, \infty[\times S$, a lower dimensional manifold without boundary. The absorbing boundary condition is given as follows.

Absorbing boundary conditions I. Approximate $z^{1/2}$ as in (5.4). Then for $u_0(x)$ given, solve the system defined by (5.5) coupled to

$$(5.6) \quad Lu = 0 \quad \text{in } [0, \infty[\times \Omega \quad u|_{t=0} = u_0 \quad \text{on } \Omega \quad \text{and}$$

$$(5.7) \quad \begin{aligned} \partial u / \partial \nu &= -\alpha(\partial_t - \Delta_S)u - \beta u - \sum_k \alpha_k (\partial_t - \Delta_S)\varphi_k \\ &\equiv B_1 u \quad \text{on } [0, \infty[\times S. \end{aligned}$$

The second family of absorbing conditions approximates $(i\tau + |\xi'|^2)^{1/2}$ differently. This is the path proposed by P. Joly [10] for constant coefficient problems in a half space. For $\xi' \neq 0$ one factors

$$(i\tau + |\xi'|^2)^{1/2} = |\xi'| (1 + i\tau/|\xi'|^2)^{1/2}.$$

Using an approximation of type (5.4) yields

$$(5.8) \quad \begin{aligned} f(z) &\equiv (1+z)^{1/2} \approx \alpha(1+z) + \beta + \sum_{k=1}^m \frac{\alpha_k(1+z)}{(1+z)+d_k}, \\ \beta &\geq 0, \quad \alpha_k > 0, \quad d_k > 0. \end{aligned}$$

Plugging in $z = i\tau/|\xi'|^2$ yields

$$(1 + i\tau/|\xi'|^2)^{1/2} \approx \alpha|\xi'|^{-2}(i\tau + |\xi'|^2) + \beta + \sum \frac{\alpha_k(|\xi'|^2 + i\tau)}{i\tau + (1 + d_k)|\xi'|^2}.$$

Multiply by $|\xi'|$ to obtain

$$(i\tau + |\xi'|^2)^{1/2} \approx |\xi'|^{-1} \left(\alpha(i\tau + |\xi'|^2) + |\xi'|^2 \left(\beta + \sum \frac{\alpha_k(|\xi'|^2 + i\tau)}{i\tau + (1 + d_k)|\xi'|^2} \right) \right).$$

This generates the approximation

$$(5.9) \quad (\partial_t - \Delta_S)^{1/2} \approx (-\Delta_S)^{-1/2} \times \left[\alpha(\partial_t - \Delta_S) - \Delta_S \left(\beta + \sum_k \alpha_k (\partial_t - \Delta_S) (\partial_t - (1 + d_k) \Delta_S)^{-1} \right) \right].$$

Introduce the auxiliary functions $\psi_k(t, x)$ on $\mathbb{R} \times S$ solutions of the parabolic initial value problems

$$(5.10) \quad (\partial_t - (1 + d_k) \Delta_S) \psi_k = u, \quad \psi_k|_{t=0} = 0, \quad k = 1, 2, \dots, m.$$

The boundary condition (5.2) is then approximated by

$$(5.11) \quad \partial u / \partial \nu = -(-\Delta_S)^{-1/2} \left[\alpha(\partial_t - \Delta_S) u - \Delta_S \left(\beta u + \sum_k \alpha_k (\partial_t - \Delta_S) \psi_k \right) \right] \equiv B_{\text{II}} u.$$

Absorbing boundary conditions II. Approximate $(1 + z)^{1/2}$ as in (5.8) then for $u_0(x)$ given solve the system defined by (5.6), (5.10), and (5.11).

In order to see that the conditions are at least well formulated, use the following simple continuity estimates for the approximations to $N_{\bar{\Omega}}$ which are immediate consequences of elementary regularity for the parabolic equations (5.12) and (5.13).

Lemma 5.1. *The operator B_{I} defined by*

$$B_{\text{I}} h \equiv -\alpha(\partial_t - \Delta_S) h - \beta h - \sum_k \alpha_k (\partial_t - \Delta_S) \varphi_k$$

where the φ_k solve

$$(5.12) \quad (\partial_t - \Delta_S + d_k) \varphi_k = h \quad \text{for } t \geq 0, \quad \varphi_k|_{t=0} = 0,$$

is a causal and bounded operator from $H^s([0, T] \times S)$ to $H^{s-2}([0, T] \times S)$ for all $s \in \mathbb{R}$.

Similarly the operator B_{II} defined by

$$B_{\text{II}} h \equiv -(-\Delta_S)^{-1/2} \left(\alpha(\partial_t - \Delta_S) h - \Delta_S \left(\beta h + \sum_k \alpha_k (\partial_t - \Delta_S) \psi_k \right) \right)$$

where the ψ_k are the solutions of

$$(5.13) \quad (\partial_t - (1 + d_k) \Delta_S) \psi_k = h, \quad \psi_k|_{t=0} = 0,$$

defines a causal and bounded operator from $H^s([0, T] \times S)$ to $H^{s-1}([0, T] \times S)$ for all $s \in \mathbb{R}$.

The absorbing boundary conditions take the form $\partial u / \partial \nu_{\Omega} = B u$ with $B = B_{\text{I}}$ or $B = B_{\text{II}}$.

The boundary operator will be applied to solutions of $Lu = 0$ which satisfy $u \in L^2([0, T] : H^1(\Omega))$. Considering u as a function of $0 \leq x_n < \varepsilon$ in geodesic normal coordinates, one has

$$u \in H^1(x_n : L^2([0, T] \times \mathbb{R}^{n-1}))$$

$$\partial u / \partial x_n \in L^2(x_n : L^2([0, T] \times \mathbb{R}^{n-1})) .$$

It follows that the trace of u at $x_n = 0$ satisfies

$$u|_{x_n=0} \in L^2([0, T] \times \mathbb{R}^{n-1}) .$$

In particular, Bu has a well defined sense for such solutions.

The absorbing strategies define well posed evolution equations for u, φ and u, ψ respectively. We give a typically parabolic well-posedness result. Estimates for higher derivatives are left to the reader.

Theorem 5.2. *For any $u_0 \in L^2(\Omega)$ and $T > 0$, there is a unique*

$$u \in C([0, T] : L^2(\Omega)) \cap H^1([0, T] : H^{-1}(\Omega)) \cap L^2([0, T] : H^1(\Omega))$$

$$\varphi_k \in C([0, T] : L^2(S)) \cap H^1([0, T] : H^{-1}(S)) \cap L^2([0, T] : H^1(S))$$

satisfying (5.5), (5.6), (5.7).

The same conclusions are valid if φ is replaced by ψ and (5.5) and (5.7) are replaced by (5.10) and (5.11).

Proof. Multiply the equation $Lu = 0$ by u and integrate $d\text{vol}_g$ over Ω . Two of the terms simplify. First,

$$(5.14) \quad - \int_{\Omega} u \Delta_g u \, d\text{vol} = \int_{\Omega} |\nabla_g u|^2 \, d\text{vol} - \int_S u \partial_{\nu} u \, d\sigma .$$

Second, note that the volume element is smoothly time dependent so

$$\int uu_t \, d\text{vol} = \partial_t \int u^2 / 2 \, d\text{vol} + \int f(t, x) u^2 \, d\text{vol}$$

where $D_{t,x}^{\alpha} f$ are bounded on $[0, \infty[\times \bar{\Omega}$. Summing we find

$$(5.15) \quad \frac{1}{2} \frac{d}{dt} \|u(t)\|_{L^2(\Omega)}^2 + \int_{\Omega} |\nabla u|^2 \, d\text{vol} - \int_{\partial\Omega} u \partial_{\nu} u \, d\sigma$$

$$+ \int_{\Omega} \left(u \sum_{j \geq 1} v_j \partial_j u + s u^2 \right) \, d\text{vol} \leq c \|u(t)\|_{L^2(\Omega)}^2 .$$

Since $v(t, x)$ and $s(t, x)$ are bounded the last integral on the left is dominated by

$$(5.16) \quad \varepsilon \|\nabla u(t)\|_{L^2(\Omega)}^2 + c \varepsilon^{-1} \|u(t)\|_{L^2(\Omega)}^2 .$$

The first of these is absorbed in the gradient term on the left and the second is passed to the right. The key is to analyse the boundary term $-\int u \partial_{\nu} u \, d\sigma$ on the left of (5.15) when $\partial_{\nu} u = Bu$.

We derive *a priori* estimates from which the existence and uniqueness results follow easily. Consider the case of B_1 first. Then

$$\begin{aligned}
-\int u \partial_\nu u d\sigma &= -\int u B_1 u d\sigma = \alpha \int u (\partial_t - \Delta_S) u d\sigma + \int \beta u^2 d\sigma \\
&\quad + \sum_k \alpha_k \int u (\partial_t - \Delta_S) \varphi_k d\sigma.
\end{aligned}$$

In the first term on the right use

$$\begin{aligned}
\int u (\partial_t - \Delta_S) u d\sigma &= \partial_t \int u^2 / 2 d\sigma + \int |\nabla_S u|^2 d\sigma \\
&\quad + \int f(t, x) u^2 d\sigma,
\end{aligned}$$

where this new f satisfies $D^\gamma f \in L^\infty([0, \infty[\times S)$ because of the smoothness of the surface element $d\sigma$. Use (5.5) in the final terms to write

$$\begin{aligned}
\int u (\partial_t - \Delta_S) \varphi_k d\sigma &= \int [(\partial_t - \Delta_S) \varphi_k]^2 d\sigma + d_k \int \varphi_k (\partial_t - \Delta_S) \varphi_k d\sigma \\
\int \varphi_k (\partial_t - \Delta_S) \varphi_k d\sigma &= \partial_t \int \varphi_k^2 / 2 d\sigma + \int |\nabla_S \varphi_k|^2 d\sigma + \int f(t, x) \varphi_k^2 d\sigma.
\end{aligned}$$

Multiplying by α and α_k and summing yields

$$\begin{aligned}
&\frac{1}{2} \frac{d}{dt} \left[\|u(t)\|_{L^2(\Omega)}^2 + \alpha \|u(t)\|_{L^2(S)}^2 + \sum \alpha_k d_k \|\varphi_k(t)\|_{L^2(S)}^2 \right] \\
&\quad + (1 - \varepsilon) \|\nabla u\|_{L^2(\Omega)}^2 + \alpha \|\nabla_S u\|_{L^2(S)}^2 + \beta \|u\|_{L^2(S)}^2 + \sum \alpha_k d_k \|\nabla_S \varphi_k\|_{L^2(S)}^2 \\
&\quad + \sum \alpha_k d_k \|(\partial_t - \Delta_S) \varphi_k\|_{L^2(\Omega)}^2 \\
(5.17) \quad &\leq c \left(\|u(t)\|_{L^2(\Omega)}^2 + \|u(t)\|_{L^2(S)}^2 + \sum \|\varphi_k(t)\|_{L^2(S)}^2 \right).
\end{aligned}$$

If $\alpha \neq 0$, Gronwall's method yields

$$\begin{aligned}
&\frac{1}{2} \left[\|u(t)\|_{L^2(\Omega)}^2 + \alpha \|u(t)\|_{L^2(S)}^2 + \sum \alpha_k d_k \|\varphi_k(t)\|_{L^2(S)}^2 \right] \\
&\quad + \int_0^t \left[(1 - \varepsilon) \|\nabla u\|_{L^2(\Omega)}^2 + \alpha \|\nabla_S u\|_{L^2(S)}^2 + \beta \|u\|_{L^2(S)}^2 \right. \\
&\quad \left. + \sum \alpha_k d_k \left[\|\nabla \varphi_k\|_{L^2(S)}^2 + \|(\partial_t - \Delta_S) \varphi_k\|_{L^2(\Omega)}^2 \right] \right] dt \\
(5.18) \quad &\leq e^{ct} \|u(0)\|_{L^2(\Omega)}^2.
\end{aligned}$$

If $\alpha = 0$, the $u|_S$ term on the right of (5.17) is first estimated by

$$\|u(t)\|_{L^2(S)}^2 \leq c \left(\eta^{-1} \|u(t)\|_{L^2(\Omega)}^2 + \eta \|\nabla u(t)\|_{L^2(\Omega)}^2 \right)$$

and then use Gronwall.

Using the equation $Lu = 0$ yields estimates for u_t in $L^2([0, T] : H^{-1}(\Omega))$.

Turn next to the case of B_{II} . The boundary integral on the left hand side of (5.15) is equal to

$$\begin{aligned}
& - \int u \partial_\nu u \, d\sigma = - \int u B_{\mathbb{1}} u \, d\sigma \\
& = \int \left[\alpha u (-\Delta_S)^{-1/2} (\partial_t - \Delta_S) u + u (-\Delta_S)^{1/2} \right. \\
& \quad \left. \times \left\{ \beta u + \sum \alpha_k (\partial_t - \Delta_S) \psi_k \right\} \right] d\sigma .
\end{aligned}$$

The terms on the right are analysed as follows.

$$\begin{aligned}
& \int u (-\Delta_S)^{-1/2} \partial_t u \, d\sigma = \partial_t \| (-\Delta_S)^{-1/4} u \|_{L^2(S)}^2 / 2 + \int f(t, x) [(-\Delta_S)^{-1/4} u]^2 \, d\sigma , \\
& \int u (-\Delta_S)^{-1/2} (-\Delta_S) u \, d\sigma = \int u (-\Delta_S)^{1/2} u \, d\sigma = \| (-\Delta_S)^{1/4} u \|_{L^2(S)}^2 .
\end{aligned}$$

The last identity is also used for the $\beta \int u (-\Delta_S)^{1/2} u \, d\sigma$ term.

$$\begin{aligned}
& \int u (-\Delta_S)^{1/2} (\partial_t - \Delta_S) \psi_k \, d\sigma = \int (\partial_t - (1 + d_k) \Delta_S) \psi_k (-\Delta_S)^{1/2} (\partial_t - \Delta_S) \psi_k \, d\sigma \\
& = \int (\partial_t - \Delta_S) \psi_k (-\Delta_S)^{1/2} (\partial_t - \Delta_S) \psi_k \, d\sigma \\
& \quad + \int -d_k \Delta_S \psi_k (-\Delta_S)^{1/2} (\partial_t - \Delta_S) \psi_k \, d\sigma \\
& = \| (-\Delta_S)^{1/4} (\partial_t - \Delta_S) \psi_k \|_{L^2(S)}^2 + d_k \| (-\Delta_S)^{5/4} \psi_k \|_{L^2(S)}^2 \\
& \quad + (d_k/2) \partial_t \| (-\Delta_S)^{3/4} \psi_k \|_{L^2(S)}^2 + o(\| (-\Delta_S)^{3/4} \psi_k \|_{L^2(S)}^2) .
\end{aligned}$$

This yields an estimate

$$\begin{aligned}
& \frac{1}{2} \left[\| u(t) \|_{L^2(\Omega)}^2 + \alpha \| (-\Delta_S)^{-1/4} u \|_{L^2(S)}^2 + \alpha_k d_k \| (-\Delta_S)^{3/4} \psi_k(t) \|_{L^2(S)}^2 \right] \\
& \quad + \int_0^t \left[\| \nabla u(t) \|_{L^2(\Omega)}^2 + \alpha \| \nabla_S u(t) \|_{L^2(S)}^2 + (\alpha + \beta) \| (-\Delta_S)^{1/4} u(t) \|_{L^2(S)}^2 \right. \\
& \quad + \sum \alpha_k \| (-\Delta_S)^{1/4} (\partial_t - \Delta_S) \psi_k(t) \|_{L^2(S)}^2 \\
& \quad \left. + \sum \alpha_k d_k \| (-\Delta_S)^{5/4} \psi_k(t) \|_{L^2(S)}^2 \right] dt \\
(5.19) & \leq e^{ct} \| u(0) \|_{L^2(\Omega)}^2 . \quad \square
\end{aligned}$$

Remarks. 1. In case $L = \partial_t - \Delta_g$ with coefficients independent of time the above energy estimates are valid with $c = 0$. The resulting stability results are uniform in $t \geq 0$ in contrast to the general case where the bounds grow exponentially in time.

2. The estimates for the functions ψ are stronger than those for φ . We will not state the corresponding strengthening in Theorem 5.2.

3. A computation like the proof of energy estimates by this integration by parts allows one to give a variational formulation of the boundary value problem which can be used as the basis of a constructive existence proof by Galerkin's method

and as the foundation of finite element methods for the discretization of the absorbing boundary value problems. In this way high order stable discretizations are available.

6. Second order absorbing conditions

In the last section, absorbing conditions were constructed using only the principal symbol of the first order operator $N_{\tilde{\Omega}}$. In this section we show how these can be refined to take advantage of the terms of order zero in the symbol of $N_{\tilde{\Omega}}$. Again, depending on the strategy for the rational approximation of $(i\tau + |\xi'|^2)^{1/2}$, there are two distinct approaches. The estimates of this section show that both approaches yield stable approximations.

In the first approach, the approximation (5.4) generated

$$(6.1) \quad (\partial_t - \Delta_S)^{1/2} \approx \alpha(\partial_t - \Delta_S) + \beta + \sum_k \alpha_k (\partial_t - \Delta_S)(\partial_t - \Delta_S + d_k)^{-1}$$

We do not improve the approximation of the first term in (4.18). Rather, the other terms are added to the algorithm. The second term in (4.18) is merely multiplication by a smooth function so causes no problem. The action of the last term on u is approximated by $\Pi(D', D')\varphi_0/2$ where φ_0 is the solution of

$$(6.2) \quad (\partial_t - \Delta_S)\varphi_0 = u, \quad \varphi_0|_{t=0} = 0.$$

For the $(1/2)V_{\tan}(\partial_t - \Delta_S)^{-1/2}$ term in (4.18) use (6.1) to get

$$(6.3) \quad \begin{aligned} (\partial_t - \Delta_S)^{-1/2} &= (\partial_t - \Delta_S)^{-1}(\partial_t - \Delta_S)^{1/2} \\ &\approx \alpha + \beta(\partial_t - \Delta_S)^{-1} + \sum_k \alpha_k (\partial_t - \Delta_S + d_k)^{-1}. \end{aligned}$$

The V_{\tan} term is then approximated by

$$(6.4) \quad V_{\tan}(\partial_t - \Delta_S)^{-1/2}u/2 \approx V_{\tan} \left[\alpha u + \beta\varphi_0 + \sum_k \alpha_k \varphi_k \right] /2$$

where the auxiliary functions φ_k are defined in (5.5). This yields the following algorithm.

Absorbing boundary condition III. Approximate $z^{1/2}$ as in (5.4) then for $u_0(x)$ given solve the coupled system for u, φ_k $0 \leq k \leq m$ defined by (5.5), (5.6), (6.2), and

$$(6.5) \quad \begin{aligned} \partial u / \partial \nu &= -\alpha(\partial_t - \Delta_S)u - \beta u - \sum_{k \geq 1} \alpha_k (\partial_t - \Delta_S)\varphi_k \\ &\quad - \frac{(n-1)H_\Omega - \langle V - \partial/\partial t, \nu_\Omega \rangle}{2} u - V_{\tan} \left[\alpha u + \beta\varphi_0 + \sum_{k \geq 1} \alpha_k \varphi_k \right] /2 \\ &\quad + \Pi(D', D')\varphi_0/2 \equiv B_{\text{III}}u \quad \text{on} \quad [0, \infty[\times S. \end{aligned}$$

Warning on signs. Theorem 4.1 gives the symbol for the Neumann operator N_Ω to Ω . The right hand side of (6.5) is an approximation to $-N_{\tilde{\Omega}}$. This explains one change of sign relative to (4.18). In addition when (4.18) is applied to $\tilde{\Omega}$, the mean curvature and outward normal are those of $\tilde{\Omega}$. In (6.5) we have chosen the normal to Ω and the mean curvature measured with that normal. For example, if Ω is a ball of radius R then $H_\Omega = 1/R$. The mean curvature and normal with respect to Ω differ by a minus sign from those of $\tilde{\Omega}$.

The second strategy of the last section is based on the approximation

$$(6.6) \quad (\partial_t - \Delta_S)^{1/2} \\ \approx (-\Delta_S)^{-1/2} \left[\alpha(\partial_t - \Delta_S) - \Delta_S \left(\beta + \sum_k \alpha_k (\partial_t - \Delta_S) (\partial_t - (1 + d_k) \Delta_S)^{-1} \right) \right].$$

As in (6.3) we then have

$$(\partial_t - \Delta_S)^{-1/2} \\ \approx (-\Delta_S)^{-1/2} \left[\alpha - \Delta_S \left(\beta (\partial_t - \Delta_S)^{-1} + \sum_k \alpha_k (\partial_t - (1 + d_k) \Delta_S)^{-1} \right) \right].$$

Use this with ψ_k defined in (5.10) to generate the approximation

$$(6.7) \quad V_{\tan} (\partial_t - \Delta_S)^{-1/2} u \\ \approx V_{\tan} (-\Delta_S)^{-1/2} \left[\alpha u - \Delta_S \left(\beta \varphi_0 + \sum \alpha_k \psi_k \right) \right]$$

Absorbing condition IV. Approximate $(1+z)^{1/2}$ as in (5.8) then for $u_0(x)$ given solve the coupled system for u, φ_0, ψ_k $1 \leq k \leq m$ defined by (5.6), (5.10), (6.2), with boundary condition

$$(6.8) \quad \partial u / \partial \nu = -(-\Delta_S)^{-1/2} \left[\alpha(\partial_t - \Delta_S)u - \Delta_S \left(\beta u + \sum \alpha_k (\partial_t - \Delta_S) \psi_k \right) \right] \\ - \frac{(n-1)H_\Omega - \langle V - \partial/\partial t, \nu_\Omega \rangle}{2} u - V_{\tan} (-\Delta_S)^{-1/2} \\ \times \left[\alpha u - \Delta_S \left(\beta \varphi_0 + \sum \alpha_k \psi_k \right) \right] / 2 + \Pi(D', D') \varphi_0 / 2 \equiv B_{IV} u$$

on $[0, \infty[\times S$.

As in the last section, we must show that these boundary value problems are well posed. The operators B_{III} and B_{IV} are causal and bounded linear maps of $H^s([0, T] \times S)$ to $H^{s-2}([0, T] \times S)$ for all $s \in \mathbb{R}$ (proof omitted). The key to the well posedness is the analysis of the new terms introduced in the boundary contribution $-\int u \partial_\nu u \, d\sigma = -\int u B u \, d\sigma$ on the left of (5.15).

Theorem 6.1. *The absorbing conditions B_{III} and B_{IV} define well posed evolution equations.*

Proof. For the first order approximation estimates (5.18) and (5.19) take the form

$$(6.9) \quad e(t) + \int_0^t E(t)dt \leq e^{ct} \|u_0(x)\|_{L^2(\Omega)}^2$$

and were derived as consequences of differential inequalities

$$(6.10) \quad de/dt + E \leq ce$$

with nonnegative quadratic forms e and E .

The boundary conditions B_{III} and B_{IV} are treated as perturbations of B_{I} and B_{II} respectively. One has

$$(6.11) \quad B_{\text{III}} = B_{\text{I}} + \alpha \partial_t + C_{\text{I}}$$

$$(6.12) \quad B_{\text{IV}} = B_{\text{II}} + \partial_t (-\Delta_S)^{-1/2} \left\{ \alpha u - \Delta \left(\beta \varphi_0 + \sum \alpha_k \psi_k \right) \right\} / 2 + C_{\text{II}}$$

where the operators C are bounded and causal from $L^2([0, T] : L^2(S))$ to itself (proof omitted).

Case 1, B_{III} . The basic inequality (6.10) is modified in two ways. First there are two new terms on the left hand side, namely

$$\int_S \alpha u \partial_t u d\sigma + \int_S u C_{\text{I}} u d\sigma .$$

Second, multiply equation (6.2) by φ_0 and perform the standard integration by parts. Add the result to (6.10).

Finally, estimate

$$\left| \int_S u C u d\sigma \right| \leq c \|u(t)\|_{L^2(S)} \|Cu\|_{L^2(S)} \\ \leq c \|u(t)\|_{L^2(S)}^2 + c \|Cu(t)\|_{L^2(S)}^2 .$$

The causality and continuity of C imply that

$$\int_0^t \|Cu(t)\|_{L^2(S)}^2 dt \leq c \int_0^t \|u(t)\|_{L^2(S)}^2 dt .$$

The next effect is to obtain an inequality of the same form as (6.10) where

$$\left(\alpha \|u(t)\|_{L^2(S)}^2 + \|\varphi_0(t)\|_{L^2(S)}^2 \right) / 2$$

is added to e and $\|\nabla \varphi(t)\|_{L^2(S)}^2$ is added to E . This suffices to prove stability.

Case 2, B_{IV} . The same strategy as above works but is more complicated. The new terms on the left hand side are

$$(1/2) \int_S u \partial_t (-\Delta_S)^{-1/2} \left[\alpha u - \Delta_S \left(\beta \varphi_0 + \sum \alpha_k \psi_k \right) \right] d\sigma \\ + \int_S u C_{\text{II}} u d\sigma .$$

The αu term is equal to

$$(1/4)(d/dt)\|(-\Delta_S)^{-1/4}u\|_{L^2(S)}^2.$$

For the β term, write

$$\begin{aligned} & \int_S u \partial_t (-\Delta_S)^{-1/2} (-\Delta_S \beta \varphi_0) d\sigma \\ &= \beta \int_S (\partial_t - \Delta_S) \varphi_0 ((-\Delta_S)^{1/2} \partial_t \varphi_0) d\sigma \\ &= \beta \|\partial_t (-\Delta_S)^{1/4} \varphi_0\|_{L^2(S)}^2 + (\beta/2)(d/dt)\|(-\Delta_S)^{3/4} \varphi_0\|_{L^2(S)}^2. \end{aligned}$$

In the same way,

$$\begin{aligned} & \int_S u \partial_t (-\Delta_S)^{-1/2} (-\Delta_S) \psi_k d\sigma \\ &= \int_S (\partial_t - (1 + d_k) \Delta_S) \psi_k \partial_t (-\Delta_S)^{1/2} \psi_k d\sigma \\ &= \|\partial_t (-\Delta_S)^{1/4} \psi_k\|_{L^2(S)}^2 + (1/2)(1 + d_k)(d/dt)\|(-\Delta_S)^{3/4} \psi_k\|_{L^2(S)}^2. \end{aligned}$$

The C term is estimated exactly as in Case 1. The net result is an estimate of form (6.10) where

$$\begin{aligned} & (\beta/2)\|\partial_t (-\Delta_S)^{3/4} \varphi_0\|_{L^2(S)}^2 \\ &+ \sum (\alpha_k/2)(1 + d_k)\|\partial_t (-\Delta_S)^{1/4} \psi_k\|_{L^2(S)}^2 \end{aligned}$$

is added to E and

$$\begin{aligned} & (1/4)\|(-\Delta_S)^{-1/4}u\|_{L^2(S)}^2 + (\beta/4)\|(-\Delta_S)^{1/4}\varphi_0\|_{L^2(S)}^2 \\ &+ \sum (\alpha_k/4)\|(-\Delta_S)^{1/4}\psi_k\|_{L^2(S)}^2 \end{aligned}$$

is added to e . This suffices to prove stability. \square

7. Rational approximations to $z^{1/2}$

Our absorbing boundary conditions are constructed with the aid of approximations to the function $z^{1/2}$. The approximants must have the form (5.4). Here $z^{1/2}$ has nonnegative real part and is defined on the complex plane with the strictly negative real axis removed.

In this section we show that such approximations can be constructed by interpolation by rational functions of suitable order at a set of points symmetric with respect to the real axis. The analysis falls into two parts. The first is to show that under these conditions, the rational interpolant is uniquely determined. Then we show that under suitable restrictions on the order, the interpolant has form

(5.4) so leads to a well posed absorbing boundary condition. The interpolants of form (5.4) are called *admissible*. The main result is Theorem 7.5.

Part I. Existence and uniqueness of interpolant

Symmetry hypothesis. We suppose that z_1, z_2, \dots, z_N are N distinct complex numbers in $\mathbb{C} \setminus]-\infty, 0[$ which are closed under complex conjugation, that is for all j , z_j^* is also a member of the set. We say that such a set is *symmetric with respect to the real axis*.

Definitions. A rational function $R = P/Q$ is of *exact type* (m, n) if P and Q are relatively prime polynomials of exact degree m and n respectively. Without loss of generality we may take Q to be monic. The rational function is *real* if $R(z)$ is real whenever z is. In that case, the coefficients of P and Q are real.

Theorem 7.1. *If the z_j satisfy the symmetry hypothesis, then for any nonnegative integers m, n with $m + n + 1 = N$, there is a unique rational function $R(z)$ of exact type (m, n) such that*

$$(7.1) \quad R(z_j) = z_j^{1/2} \quad \text{for } j = 1, 2, \dots, N .$$

In addition, the interpolant is real.

Remark. The symmetry is crucial. In contrast to the case of polynomial interpolation, having the correct number of interpolation points is not sufficient. After the proof we present an example with $n = 1$, $m = 3$, and a nonsymmetric set of points for which interpolation fails.

Proof of Theorem. First we show that reality follows from uniqueness. Let $T(z)$ denote the rational function obtained from R by taking the complex conjugate of the coefficients. Then $R(z)^* = T(z)^*$ for all z . The symmetry hypothesis implies that $R(z^*) = R(z)^*$ when z is an interpolation point. Thus, R and T agree at the N points z_j and therefore must be identical.

For $j \leq N$ let $x_j \equiv z_j^{1/2}$, $\text{Re}(x_j) > 0$. We must find polynomials P, Q of degree m and n respectively with Q monic and such that

$$(7.2) \quad P(x^2) - xQ(x^2) = 0 \quad \text{at the points } x_1, x_2, \dots, x_N .$$

The condition (7.2) gives a system of N linear equations for the undetermined coefficients of the polynomials P and Q . It suffices to show that the determinant of the coefficient matrix does not vanish. The j^{th} row of the coefficient matrix is equal to

$$r_j \equiv (1, x_j^2, x_j^4, \dots, x_j^{2m}, x_j, x_j^3, \dots, x_j^{2n-1}) .$$

The determinant is denoted $D(x_1, \dots, x_N)$.

If one of the x_j 's is zero, we may renumber so that $x_1 = 0$. Expanding the determinant along the first row shows that

$$D(x_1, \dots, x_N) = x_2 \dots x_N D(x_2, \dots, x_N)$$

so it suffices to treat the cas of nonzero x_j 's.

Next permute the columns of D so that the exponents are decreasing. The first $k \equiv (|2n - 1 - 2m| - 1)/2$ exponents decrease by two. The rest decrease by one. Thus, if λ_i is the nonincreasing sequence defined by

$$(7.3) \quad \lambda_1 = k, \quad \lambda_2 = k - 1, \dots, \lambda_k = 1, \quad \lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_N = 0$$

then still denoting by r_j the rows after the permutation,

$$r_j = (x_j)^{\lambda_i - i + N} \quad i = 1, 2, \dots, N.$$

Since $x_i - x_j = 0$ implies that $D = 0$, and $x_i - x_j$ is irreducible it follows that D is divisible by $x_i - x_j$ in $\mathbb{Z}(x_1, \dots, x_N)$. Therefore, D is divisible by the Vandermonde determinant

$$\det \left[(x_j)^{N-i} \right]_{1 \leq i, j \leq N} = \prod_{1 \leq i < j \leq N} (x_i - x_j).$$

The quotient is called the *Schur determinant* associated to λ and is denoted $\chi_k(x_1, \dots, x_N)$. From the definition note that permuting two of the x_j 's does not change the value of χ so χ is a symmetric function. It suffices to prove that $\chi_k(x_1, \dots, x_n) \neq 0$. Since $\chi_0 = 1$ the case $k = 0$ is trivial. \square

Theorem 7.2. *If $\{x_1, \dots, x_N\} \subset \{\operatorname{Re}(z) > 0\}$ is a set of distinct complex numbers which is symmetric with respect to the real axis, then $\chi_k(x_1, \dots, x_N) \neq 0$.*

Proof. Denote by σ_j the elementary symmetry function of degree j , that is

$$\sigma_j(x_1, \dots, x_N) \equiv \sum_{1 \leq i_1 < i_2 < \dots < i_j \leq N} x_{i_1} \dots x_{i_j}.$$

Extend this definition by the conventions $\sigma_0 \equiv 1$ and $\sigma_j \equiv 0$ whenever $j < 0$ or $j > N$.

Order the x_j 's with the reals last as follows

$$(7.4) \quad x_1, \dots, x_I, \bar{x}_1, \dots, \bar{x}_I, x_{2I+1}, \dots, x_N.$$

Let S be the monic polynomial with the x_j as roots so

$$(7.5) \quad S(z) = \prod_{i \leq I} (z - x_i)(z - \bar{x}_i) \prod_{i > 2I} (z - x_i) = \sum_{j=0}^N (-1)^{N-j} \sigma_{N-j} z^j.$$

Since the quadratic and linear factors have real coefficients it follows that the $\sigma_j(x_1, \dots, x_N)$ are real. The quadratic factors $z^2 - 2\operatorname{Re}(x_i)z + |x_i|^2$ and linear factors both have nonvanishing coefficients of alternating signs. The same is therefore true of the product so

$$(7.6) \quad \sigma_J(x_1, \dots, x_N) > 0.$$

For a nonincreasing sequence λ_j the *dual sequence* λ' is defined by

$$\lambda'_i \equiv \#\{j | \lambda_j \geq i\}.$$

The sequence defined by (7.3) is its own dual.

Jacobi's identity [14, formula 3.5] applied to our λ and with $m = k$ reads

$$(7.7) \quad \chi_k(x_1, \dots, x_N) = \det \left[(\sigma_{\lambda'_i - i + j})_{1 \leq i, j \leq k} \right] = \det \left[(\sigma_{k - 2i + j + 1})_{1 \leq i, j \leq k} \right].$$

The last equality follows since $\lambda'_i = \lambda_i = (k - i + 1)_+$ implies $\lambda'_i - i + j = k - 2i + j + 1$ for $i \leq k$.

To show that χ is nonzero we show that the matrix appearing on the right of (7.7) is nonsingular. The matrices for $k = 1, 2, 3$ are

$$\sigma_1 \quad \begin{bmatrix} \sigma_2 & \sigma_3 \\ \sigma_0 & \sigma_1 \end{bmatrix} \quad \begin{bmatrix} \sigma_3 & \sigma_4 & \sigma_5 \\ \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_{-1} & \sigma_0 & \sigma_1 \end{bmatrix}.$$

Since $\sigma_1 = \sum x_i > 0$, the $k = 1$ case is nonsingular. Note that the lower right $(k - 1) \times (k - 1)$ minor is equal to the previous matrix and that all terms more than one below the principal diagonal vanish.

A vector $(-a_0, +a_1, \dots, (-1)^{j+1}a_j, \dots, (-1)^k a_{k-1})$ belongs to the nullspace if and only if for $1 \leq i \leq k$

$$(7.8) \quad 0 = \sum_{j=1}^k (1)^j a_{j-1} \sigma_{k-2i+j+1} = \sum_{j=0}^{k-1} (-1)^{j+1} a_j \sigma_{k-2i+j+2}.$$

Let $R(x)$ be the polynomial of degree $d \leq k - 1$ defined by $R(x) \equiv \sum a_j x^j$. With S defined in (7.5) one has

$$(7.9) \quad RS = \sum_{h=0}^{N+k-1} \left(\sum_{j=0}^{k-1} (-1)^{N-h+j} \sigma_{N-h+j} a_j \right) x^h$$

with the convention that σ_j vanishes if $j < 0$ or $j > N$. Thus (7.8) is equivalent to the vanishing of the terms of order $h = N + 2i - k - 2$ in RS for $1 \leq i \leq k$. These are k alternating terms beginning with the term of order $N + k - 2$, which is the next to leading order.

Since d denotes the degree of R , $a_d \neq 0$ and $a_j = 0$ for $j > d$. Thanks to (7.6) and the form of the matrix of σ 's, d must be greater than or equal to 1 for any element of the nullspace.

The product RS then has the form $RS = a_d x^{N+d} + \text{lot}$ so d and k must have opposite parity, otherwise (7.8), (7.9) would show that this leading term would vanish. Thus RS has the form

$$RS = a_d x^{N+d} + b_{N+d-2} x^{N+d-2} + b_{N+d-4} x^{N+d-4} + \dots \\ + b_{N-k+1} x^{N-k+1} + 0(x^{N-k}) + \text{lower order}.$$

If $N + d$ is even then the even part is of order $N + d$ and the first possible odd term in x^{N-k-2} so the even and odd parts of RS have degrees differing by at least $d + k + 2$. If $N + d$ is odd, the odd part has degree $N + d$ and the first possible even term is in x^{N-k+2} . The difference of degrees is again at least $d + k + 2$. Since $d \leq k - 1$, this shows that in both cases the even and odd parts of RS have degrees differing by at least $2d + 3$.

Theorem 7.2 is an immediate consequence of the following Lemma which is the main result of this section.

Lemma 7.3. *Suppose that $U(x)$ is a real polynomial whose even and odd parts have degrees differing by $2K + 1$. Denote by $M^+(U)$ (resp. $M^-(U)$) the number of roots of U with strictly positive (resp. negative) real parts. Then $M^+(U) \geq K$ and $M^-(U) \geq K$.*

End of proof of Theorem 7.2 assuming Lemma 7.3. Apply the lemma to the polynomial $U = RS$.

The remark before the Lemma shows that $K \geq d + 1$. However, the only roots of RS in the half plane $\text{Re}(z) < 0$ are roots of R so $M^- \leq d$. This contradiction shows that no nontrivial element of the nullspace exists. \square

Proof of Lemma. By hypothesis $U(x) \equiv P(x^2) - xQ(x^2)$ with real polynomials P and Q of exact degrees p and q respectively. The difference of degrees of the even and odd parts of U is then equal to $|2p - (2q + 1)|$. Thus K is given by $2K + 1 \equiv |2p - (2q + 1)|$.

If the polynomials P and Q are not relatively prime, then one can write $U = G\tilde{U}$ with $\tilde{U} = \tilde{P}(x^2) - x\tilde{Q}(x^2)$ with \tilde{P} and \tilde{Q} relatively prime. Then $M^\pm(U) \geq M^\pm(\tilde{U})$ and $K(U) = K(\tilde{U})$. Thus it suffices to prove the result for \tilde{U} . Thus we may suppose that P and Q are relatively prime.

If $P(0) = 0$ then $Q(0) \neq 0$ for otherwise x would be a common factor. Factor $U = x\tilde{U}$ with $\tilde{U} = \tilde{P}(x^2) - x\tilde{Q}(x^2)$ where $\tilde{P} = -Q$ and $\tilde{Q} = P/x$. Then $2\tilde{p} - 2\tilde{q} - 1 = -(2p - 2q - 1)$ so $K(\tilde{U}) = K(U)$. Thus it suffices to prove the lemma for \tilde{U} , so, we may suppose that $P(0) \neq 0$.

We split the proof according to the parity of $\deg(U)$.

Case 1. $\deg(U)$ is even, that is $2p > 2q + 1$. Then $K = p - q - 1$. The strategy is to deform Q to a polynomial of degree zero. If Q is not of degree zero, denote by $\alpha > 0$ the largest integer such that x^α divides Q . Then

$$\begin{aligned} Q(x) &= a_q x^q + \dots + a_\alpha x^\alpha, \quad a_q \neq 0 \neq a_{\alpha'} \\ U(x) &= P(x^2) - x^{2\alpha+1} Q_1(x^2), \\ Q_1(x) &\equiv a_q x^{q-\alpha} + \dots + a_\alpha. \end{aligned}$$

Let

$$\begin{aligned} Q_t(x) &\equiv t(a_q x^{q-\alpha} + \dots + a_{\alpha+1} x^{\alpha+1}) + a_\alpha, \quad \text{and} \\ U_t(x) &\equiv P(x^2) - x^{2\alpha+1} Q_t(x^2). \end{aligned}$$

If the roots of U_t are never purely imaginary then U_0 satisfies

$$M^\pm(U_0) = M^\pm(U), \quad U_0 = P(x^2) + x^{2\alpha+1}Q_0(x^2),$$

$$\deg Q_0 < \deg Q, \quad K(U_0) > K(U).$$

If on the other hand t_0 is the first $t < 1$ such that U_t has roots on the imaginary axis, write

$$U_{t_0}(x) = \prod (x^2 + \lambda_j^2) \tilde{U}(x)$$

with λ_j real and the roots of \tilde{U} with nonzero real parts. Write

$$\tilde{U}_{t_0}(x) \equiv \tilde{P}(x^2) - x^{2\alpha+1}\tilde{Q}(x^2).$$

Then

$$M^\pm(\tilde{U}) < M^\pm(U), \quad \deg(\tilde{P}) < \deg(P),$$

$$\deg(\tilde{Q}) < \deg(Q), \quad 2\tilde{p} > 2\tilde{q} + 1, \quad K(\tilde{U}) = K(U).$$

Both cases lead to a new polynomial with $2\tilde{p} > 2\tilde{q} + 1$ and with M^\pm non-increasing, K nondecreasing and the degree of Q strictly decreasing. A finite number of iterations leads to a polynomial Q of degree zero, that is

$$\tilde{U}(x) = \tilde{P}(x^2) - cx^{2\alpha+1}, \quad c \neq 0, \quad \tilde{P}(0) \neq 0,$$

$$M^\pm(\tilde{U}) \leq M^\pm(U), \quad K(\tilde{U}) \geq K(U).$$

On $i\mathbb{R}$, the first term of \tilde{U} is real and is nonzero at 0 while the second is imaginary and nonzero on $\mathbb{R} \setminus 0$. Thus, such a polynomial cannot have purely imaginary roots. A deformation leaving M^\pm and K invariant reduces to the case

$$\tilde{U}(x) = x^{2p} \pm 1 \pm x^{2\alpha+1} \quad K = p - \alpha - 1.$$

We must show that $M^\pm \geq K$. Consider the deformations

$$x^{2p} \pm 1 \pm \varepsilon x^{2\alpha+1}$$

with ε decreasing from 1 to 0. The numbers M^\pm are independent of ε . For ε equal to zero at most 2 of the $2p$ roots lie on the imaginary axis. Thus $2p - 2$ roots are off axis and remain so for ε small. At least $p - 1$ of these lie in each half space, so for ε small one has $M^\pm \geq p - 1 \geq p - \alpha - 1 = K$. This completes the proof for case 1.

Case 2. $\deg(U)$ is odd, that is $2p < 2q + 1$. The strategy is to deform P to a polynomial of degree zero. With t decreasing from 1 to zero let

$$P(x) = b_p x^p + \dots + b_1 x + b_0,$$

$$P_t(x) \equiv t(b_p x^p + \dots + b_1 x) + b_0, \quad \text{and}$$

$$U_t(x) \equiv P_t(x^2) - xQ(x^2).$$

If complex roots appear one extracts factors of $x^2 + \lambda_j^2$ as in case 1. If not, one sets t equal to zero. Either way one gets a new polynomial \tilde{U} with M^\pm nonincreasing, K nondecreasing, and the degree of P strictly decreasing. A finite number of iterations leads to

$$\tilde{U}(x) = b - x\tilde{Q}(x^2), \quad b \neq 0.$$

Such polynomials cannot have purely imaginary roots so a deformation reduces to the case

$$\tilde{U}(x) = \pm 1 \pm x^{2q+1}.$$

The reflection $x \mapsto -x$ reduces to $\pm 1 + x^{2q+1}$. In this case there are no purely imaginary roots and at least q roots in each half space. Thus $M^\pm \geq q = K$ and the proof is complete. \square

Nonsymmetric counterexample. We construct five points x_j in the half-plane $\operatorname{Re}(x) > 0$ such that there does not exist an interpolating rational function $R = P/Q$ of type (3,1) with Q monic. The interpolation (7.1) holds if and only if the coefficients $(p_3, p_2, p_1, p_0, q_1)$ satisfy the linear system whose coefficient matrix has rows equal to

$$(1, x_j^2, x_j^4, x_j^6, x_j) \quad j = 1, \dots, 5$$

and right hand side equal to (x_j^3) . In order that matrix not be surjective one must choose the points so that the determinant of the coefficient matrix vanishes, that is

$$(7.10) \quad \chi_2(x_1, \dots, x_5) = 0.$$

We will choose the points so that in addition

$$(7.11) \quad \chi_2(x_1, \dots, x_4) \neq 0$$

which implies that the last four columns of the coefficient matrix are linearly independent and therefore span the range. In that case, the vector (x_j^3) belongs to the range if and only if

$$\det(x_i^3, x_i^2, x_i^4, x_i^6, x_i) = 0 \quad \text{iff} \quad \chi_1(x_1, \dots, x_5) = \sum x_j = 0.$$

The last equality cannot hold since the x_j have positive real parts. Thus it suffices to find points satisfying (7.10), (7.11).

Our interest being $N = 5$, compute

$$\begin{aligned} \chi_2(x_1, \dots, x_N) &= \sum_{i \neq j} x_i^2 x_j + 2 \sum_{i < j < k} x_j x_j x_k = [(\sum x_i)^3 - \sum x_i^3]/3 \\ &= \left(\sum_1^{N-1} x_i \right) x_N^2 + \left(\sum_1^{N-1} x_i \right)^2 x_N + \chi_2(x_1, \dots, x_{N-1}). \end{aligned}$$

Choose with the same imaginary parts, $x_j = \alpha + \beta_j$ with

$$(7.12) \quad \sum_1^{N-1} \beta_j = 0.$$

Then

$$(7.13) \quad \begin{aligned} \chi_2(x_1, \dots, x_{N-1}) &= N(N-1)(N-2)\alpha^3/3 + \alpha \sum_1^{N-1} \beta_j^2 + i \sum_1^{N-1} \beta_j^3/3. \\ \chi_2(x_1, \dots, x_N) &= (\alpha + i\beta_N)(N-1)\alpha(N\alpha + i\beta_N) + \chi_2(x_1, \dots, x_{N-1}). \end{aligned}$$

The β_j are chosen with

$$(7.14) \quad \sum_1^{N-1} \beta_j^3 = 1$$

so $\chi_2(x_1, \dots, x_{N-1}) \neq 0$ which yields (7.11).

Equation (7.13) shows that $\chi_2(x_1, \dots, x_N) = 0$ if and only if

$$(7.15) \quad \begin{aligned} N(N^2 - 1)\alpha^2 - 3(N-1)\beta_N^2 + 3 \sum_1^{N-1} \beta_j^2 &= 0, \quad \text{and} \\ (N^2 - 1)\alpha^2\beta_N + 1/3 &= 0. \end{aligned}$$

Eliminating α^2 yields a cubic equation for β_N

$$9(N-1)\beta_N^3 - \left(9 \sum_1^{N-1} \beta_j^2\right)\beta_N + N = 0.$$

Choose $\beta_1, \dots, \beta_{N-1}$ so that $\beta_N = -1$ is a solution. That is

$$(7.16) \quad \sum_1^{N-1} \beta_j^2 = (8N-9)/9.$$

Then $\alpha > 0$ is determined by (7.15).

To finish the construction it suffices to choose $\beta_1, \dots, \beta_{N-1}$ distinct in $\mathbb{R} \setminus \{1\}$ satisfying (7.12), (7.14), (7.16) which is possible since $N-1 > 3$.

Part II. Partial fraction decomposition of the interpolant

In order that the approximant generate an artificial boundary condition which can be implemented by solving surface differential equations no worse than heat equations we restrict attention to rational approximations of the form

$$(7.17) \quad R(z) = \alpha z + \beta + \sum \alpha_i z / (z + d_i).$$

Since R grows at most linearly at infinity this restricts to rational functions of exact order (m, n) with $m \leq n + 1$. In order that the resulting artificial boundary conditions satisfy an energy identity which implies well posedness, we impose the following admissibility condition.

Definition. A rational function $R(z)$ is admissible if its partial fraction decomposition has form (7.17) with $\alpha \geq 0$, $\beta \geq 0$, $\alpha_i > 0$, and $d_i > 0$.

Note that admissibility implies reality.

Proposition 7.4. *An admissible rational function of exact type (m, n) must satisfy $n + 1 \geq m \geq n$.*

Proof. Since R grows at most linearly at infinity one must have $n + 1 \geq m$. It remains to show that $m \geq n$.

If $m \leq n - 1$, then as z tends to infinity, $R(z)$ tends to zero. However, for an admissible rational function,

$$R(z) = \alpha z + \left(\beta + \sum \alpha_i \right) + O(1/z).$$

Since the $\beta + \sum \alpha_i > 0$, this cannot tend to zero. \square

To study fractions of exact type (m, n) with $n \leq m \leq n + 1$ which interpolate $z^{1/2}$ at $N = m + n + 1$ points z_1, \dots, z_N symmetric with respect to the real axis a crucial role is played by $S \equiv S_N(z)$ defined in (7.5). As above $x_i \equiv z_i^{1/2}$ are ordered as in (7.4). Denote by $\tilde{P}(z^2)$ and $z\tilde{Q}(z^2)$ the even and odd parts of S defined by

$$\begin{aligned} \tilde{P}(z^2) &\equiv [S(z) + S(-z)]/2. \\ z\tilde{Q}(z^2) &\equiv [S(z) - S(-z)]/2. \end{aligned}$$

Then \tilde{P} and \tilde{Q} are polynomials of exact degree $\lfloor N/2 \rfloor$ and $\lfloor (N-1)/2 \rfloor$ where $\lfloor \cdot \rfloor$ denotes *floor function*, that is the largest integer less than or equal to.

We next study the negative real roots of $\tilde{P}(w)$ and $\tilde{Q}(w)$. These correspond to purely imaginary roots of the polynomials $\tilde{P}(z^2)$ and $\tilde{Q}(z^2)$. Writing

$$\frac{S(iy)}{S(-iy)} = \frac{\tilde{P}(-y^2) + iy\tilde{Q}(-y^2)}{\tilde{P}(-y^2) - iy\tilde{Q}(-y^2)}$$

shows that this is equivalent to studying the real roots of $S(iy)/S(-iy) = \pm 1$ respectively.

Thanks to the symmetry of the x_k one has

$$S(iy)/S(-iy) = \prod (iy - x_k) / \prod (-iy - x_k) = \prod (iy - x_k) / \prod (-iy - \bar{x}_k).$$

This is a product of N factors of modulus 1. We study its argument as y increases from 0 to ∞ . Each of the factors is

$$(x_k - iy) / (\bar{x}_k + iy) = e^{2i\theta_k(y)}$$

where the argument $\theta_k(y)$ decreases from $\arg(x_k)$ to $-\pi/2$ as y increases from 0 to ∞ . The branch of the argument is defined by $-\pi/2 < \arg < \pi/2$. In particular $\arg(x_k) + \arg(\bar{x}_k) = 0$. Thus

$$S(iy)/S(-iy) = e^{i\theta(y)} \quad \text{where} \quad \theta(y) \equiv 2 \sum \theta_k(y),$$

and θ decreases from 0 to $-N\pi$ as y increases from 0 to ∞ .

$S(iy)/S(-iy)$ is equal to $+1$ (resp. -1) when θ is equal to $-n\pi$ with n an odd (resp. even) integer in $]0, N[$. This gives $\lfloor N/2 \rfloor$ (resp. $\lfloor (N-1)/2 \rfloor$) values

of y such that $S(iy)/S(-iy) = -1$ (resp. $+1$). Denote the values as $y_1, \dots, y_{\lfloor N/2 \rfloor}$ (resp. $Y_1, \dots, Y_{\lfloor (N-1)/2 \rfloor}$). The roots are interlaced so that

$$(7.18) \quad y_1 < Y_1 < y_2 < \dots < y_{\lfloor N/2 \rfloor - 1} < Y_{\lfloor N/2 \rfloor - 1} < y_{N/2} \text{ if } N \text{ is even}$$

$$(7.19) \quad y_1 < Y_1 < y_2 < \dots < y_{(N-1)/2} < Y_{(N-1)/2} \text{ if } N \text{ is odd .}$$

This gives exactly $\lfloor N/2 \rfloor$ negative real zeros, $-y_j^2$, of \tilde{P} and $\lfloor (N-1)/2 \rfloor$ negative real zeros, $-Y_j^2$, of \tilde{Q} with a root of \tilde{P} closest to the origin. The number of zeros is equal to the order of the corresponding polynomial in all cases.

The signs of $\tilde{P}(Y_j^2)$ (resp. $\tilde{Q}(y_j^2)$) alternate. The signs are determined by the signs of the extreme elements which in turn are determined by the asymptotics on the real axis.

The leading terms in \tilde{P} and \tilde{Q} in case $N = 2p$ is even are

$$(7.20) \quad \tilde{P}(w) = w^p + \text{lower order} , \quad \tilde{Q}(w) = -\sigma_1 w^{p-1} + \text{lower order} ,$$

where as before $\sigma_1 \equiv \sum x_j$. On $] - y_1^2, \infty[$, \tilde{P} is positive. The alternation then shows that

$$(7.21) \quad (-1)^j \tilde{P}(-Y_j^2) > 0 \text{ when } N \text{ is even.}$$

If $N = 2p + 1$ is odd, then

$$(7.22) \quad \tilde{P}(w) = -\sigma_1 w^p + \text{lower order} , \quad \tilde{Q}(w) = w^p + \text{lower order.}$$

Reasoning as above shows that $\tilde{P} < 0$ on $] - y_1^2, \infty[$ and

$$(7.23) \quad (-1)^{j+1} \tilde{P}(-Y_j^2) > 0 .$$

Theorem 7.5. *Suppose that z_1, \dots, z_N are distinct points in $\mathbb{C} \setminus]0, \infty[$ symmetric with respect to the real axis and containing at least one nonzero point. Then, there is exactly one admissible rational function interpolating $z^{1/2}$ at these points. That rational function is real and of type (m, m) when $N = 2m + 1$ is odd and of type $(m, m - 1)$ when $N = 2m$ is even. In both cases the interpolant is given by the formula $R(z) = -\tilde{P}(z)/\tilde{Q}(z)$ where $\tilde{P}(z^2)$ and $z\tilde{Q}(z^2)$ are the even and odd parts of $S(z) \equiv \prod (z - z_j^{1/2})$.*

Proof. The proof is a case by case analysis of rational interpolants of type (m, n) with $n \leq m \leq n + 1$.

We seek P and monic Q of exact degrees m and n respectively such that $P(x^2) - xQ(x^2)$ vanishes at the points x_1, \dots, x_N . Recall that $N = m + n + 1$ and $x_j = z_j^{1/2}$. Thus $S(x)$ divides $P(x^2) - xQ(x^2)$.

Case 1, $m = n$. Then $N = 2n + 1$ is odd and $-P(x^2) + xQ(x^2)$ is a monic polynomial of degree $2n + 1$ which divides S which is also monic and of degree $2n + 1$. It follows that the two polynomials must be equal.

Thus $-P$ and xQ are the even and odd parts of S from the statement of the theorem,

$$P(z) = -\tilde{P}(z), \quad Q(z) = \tilde{Q}(z).$$

Q is monic and the leading term of P is $\sigma_1 z^n$. $\operatorname{Re}(\sigma_1) \geq 0$ since $\operatorname{Re}(x_j) \geq 0$. In fact the real part of σ_1 is strictly positive unless the set of z_j is the singleton $\{0\}$ which is ruled out by hypothesis. Thus the exact degree is (n, n) .

To prove admissibility, decompose R into partial fractions. We have observed that Q has n distinct negative real roots q_j , so

$$P(z)/Q(z) \sim \sigma_1 + \sum \rho_j/(z - q_j).$$

Rewrite this in the form

$$P(z)/Q(z) \sim \beta + \sum \alpha_j z/(z - q_j).$$

Then

$$\beta = P(0)/Q(0) = -\tilde{P}(0)/\tilde{Q}(0) = \sigma_{2n+1}/\sigma_{2n}.$$

the last equality following from (7.5) and the relation between \tilde{P}, \tilde{Q} and the even and odd parts of S . Compute

$$\begin{aligned} \sigma_{2n+1} &= \prod x_j = \prod_{j \leq l} |x_j|^2 \prod_{j > 2l} x_j > 0, \quad \text{and} \\ \sigma_{2n} &= \sum_{i=1}^N \left(\prod_{j \neq i} x_j \right) = \sigma_{2n+1} \sum 1/x_j > 0. \end{aligned}$$

Thus, $\beta > 0$.

To find the sign of α_j use the formula

$$q_j \alpha_j = -\tilde{P}(q_j) / \prod_{i \neq j} > (q_j - q_i).$$

Then (7.16) implies that $\alpha_j > 0$.

Thus $R = P/Q$ is admissible.

Case 2, $m = n + 1$. In this case $N = 2m$ is even and (7.20) shows that $P(x^2) - xQ(x^2)$ is a monic polynomial of degree $2m$. As this polynomial divides S which is also monic of degree $2m$, the two polynomials must be equal. Thus we must have

$$P(z) = \tilde{P}(z), \quad \text{and} \quad Q(z) = -\tilde{Q}(z).$$

Decompose as in case 1,

$$P(z)/Q(z) = \alpha z + \beta + \sum \alpha_j z/(z - q_j).$$

Then

$$\begin{aligned} \alpha &= 1/\sigma_1 > 0, \\ \beta &= -\tilde{P}(0)/\tilde{Q}(0) = \sigma_{2n}/\sigma_{2n-1} > 0, \quad \text{and} \\ q_j \alpha_j &= -\tilde{P}(q_j) / \left(\sigma_1 \prod_{i \neq j} (q_j - q_i) \right) < 0, \end{aligned}$$

the last inequality relying on (7.23). It follows that $\alpha_j > 0$ and we conclude that R is admissible. \square

8. Estimates of the truncation error

In Theorems 5.2 and 6.1 it was shown that the absorbing boundary conditions define parabolic time evolutions. This suggests that the error $v \equiv u_{\text{ex}} - u_{\text{apprx}}$ be considered as a solution of the boundary value problem

$$(8.1) \quad Lv = 0 \text{ in } \mathbb{R} \times \Omega, \quad v = 0 \text{ for } t \leq 0, \quad \partial v / \partial \nu = Bv + (B + N_{\tilde{\Omega}})u_{\text{ex}}$$

where the equation $\partial u_{\text{ex}} / \partial \nu = -N_{\tilde{\Omega}}u_{\text{ex}}$ for the exact solution u_{ex} is used in the last equality. Recall that B is constructed as an approximation to $-N_{\tilde{\Omega}}$ so it is expected that $(B + N_{\tilde{\Omega}})u_{\text{ex}}$ will be small in some sense. The size of the error is bounded by the size of this quantity.

Theorem 8.1. *There is a constant c , independent of the absorbing boundary condition $\partial u_{\text{apprx}} / \partial \nu = Bu$ constructed in the last sections, such that*

$$(8.2) \quad \begin{aligned} & \| (u_{\text{ex}} - u_{\text{apprx}})(t) \|_{L^2(\Omega)}^2 + \int_0^t \| \nabla (u_{\text{ex}} - u_{\text{apprx}})(t) \|_{L^2(\Omega)}^2 dt \\ & \leq e^{ct} \int_0^t \| (B + N_{\tilde{\Omega}})u_{\text{ex}} \|_{L^2(S)}^2 dt . \end{aligned}$$

Proof. Let $v \equiv u_{\text{ex}} - u_{\text{apprx}}$,

$$\begin{aligned} e(t) & \equiv \| (u_{\text{ex}} - u_{\text{apprx}})(t) \|_{L^2(\Omega)}^2, \\ E(t) & \equiv \| \nabla (u_{\text{ex}} - u_{\text{apprx}})(t) \|_{L^2(\Omega)}^2. \end{aligned}$$

The standard parabolic estimate for v then reads

$$(8.3) \quad de/dt + 2E \leq ce + \int_S vBvd\sigma + \int_S v(B - N_{\tilde{\Omega}})u_{\text{ex}}d\sigma .$$

The derivation of the stability estimates for B shows that

$$\int_0^t \int_S vBvd\sigma dt \leq \varepsilon \int_0^t E(t)dt + c\varepsilon^{-1} \int_0^t e(t)dt .$$

The integral of the last term in (8.3) is estimated as follows.

$$\begin{aligned} & \int_0^t \int_S v(B - N_{\tilde{\Omega}})u_{\text{ex}}d\sigma dt \leq c \int_0^t \| v(t) \|_{L^2(S)}^2 dt \\ & \quad + \int_0^t \| (B - N_{\tilde{\Omega}})u_{\text{ex}}(t) \|_{L^2(S)}^2 dt \\ & \leq \varepsilon \int_0^t E(t)dt + c\varepsilon^{-1} \int_0^t e(t)dt + \int_0^t \| (B - N_{\tilde{\Omega}})u_{\text{ex}}(t) \|_{L^2(S)}^2 dt . \end{aligned}$$

Therefore

$$(8.4) \quad e(t) + \int_0^t E(t)dt \leq c \int_0^t e(t)dt + \int_0^t \|(B - N_{\bar{\Omega}})u_{\text{ex}}(t)\|_{L^2(S)}^2 dt$$

with constant c uniform for the absorbing conditions constructed in the previous sections.

Gronwall's method applied to (8.4) yields (8.2). \square

There is a second estimate which is relevant for our absorbing conditions. This one is qualitative in nature. The boundary conditions are constructed by approximating the symbol of the Neumann operator $N_{\bar{\Omega}}$ by suitable rational functions. Only the terms of order zero and one in the symbol are used as these are given in terms of simple geometric quantities associated with S . The error has two sources. First there is the error committed by taking only one or two terms from the symbol which approximates the Neumann operator by N_{apprx} which differs from N by an element of OpS_{Heat}^m with $m = 0$ or $m = -1$, depending on whether the first order or second order approach is used. Second, there is the error from approximating N_{apprx} by an operator B which has symbol an appropriate rational function. The next theorem considers the error upon replacing N by N_{apprx} . The result is that the error is smoother than u_{ex} near to $t = 0$. If $N - N_{\text{apprx}}$ is in OpS^{-1} the degree of smoothness is higher than if $N - N_{\text{apprx}}$ is in OpS^0 .

Suppose that

$$(8.5) \quad N_{\text{apprx}} \text{ is a causal member of } OpS_{\text{Heat}}^1(\mathbb{R} \times S), \text{ and}$$

$$(8.6) \quad N_{\bar{\Omega}} - N_{\text{apprx}} \in OpS_{\text{Heat}}^m(\mathbb{R} \times S), \quad m = 0 \quad \text{or} \quad m = -1.$$

For $u_0 \in H^m(\mathbb{R}^n)$ with $\text{supp } u_0 \subset \bar{\Omega}$ let u_{ex} denote the solution of $Lu_{\text{ex}} = 0$ on $[0, \infty[\times \mathbb{R}^n$ with initial data u_0 . Let u denote the solution in $[0, \infty[\times \Omega$ defined by the same initial data and the absorbing boundary condition $\partial u / \partial \nu = -N_{\text{apprx}}u$. It is not difficult to show that this boundary condition defines a stable parabolic evolution.

The error $w \equiv u_{\text{ex}} - u$, extended by zero for $t < 0$, is then a solution of

$$(8.7) \quad Lw = 0 \quad \text{in }] - \infty, \infty[\times \Omega, \quad w = 0 \quad \text{in } t \leq 0,$$

$$(8.8) \quad \partial_\nu w = -N_{\text{apprx}}w - (N_{\bar{\Omega}} - N_{\text{apprx}})u_{\text{ex}} \quad \text{on } [0, \infty[\times S.$$

The idea for estimating w is that the difference of the N 's belongs to OpS_{Heat}^m so is a lower order term. Whereas both u_{ex} and u are continuous with values in $H^m(\Omega)$, the difference is continuous with values in $H^1(\Omega)$. In this sense u gives a good approximation of the singular and most interesting behavior of u_{ex} .

Theorem 8.2. *If N_{apprx} satisfies (8.5) and (8.6) then, the error w is continuous on $[0, \infty[$ with values in $H^m(\Omega)$. In addition there is a constant $c(T)$ such that*

$$(8.9) \quad \|w(t)\|_{H^1(\Omega)} \leq c \|u_0\|_{H^m(\Omega)} \quad 0 \leq t \leq T.$$

The regularizing property of the heat equation implies that the $H^1(\Omega)$ norm of both u_{ex} and of u are finite for $t > 0$ but they usually grow like t^{-1+m} times the

$H^m(\Omega)$ norm of u_0 as t decreases to zero. Thus (8.9) shows that u is genuinely an approximation to u_{ex} for t small.

Proof. We derive (8.9) for regular solutions. Standard techniques suffice from there.

Multiply the equation $Lw = 0$ by w and integrate over Ω to find that

$$(8.10) \quad \frac{d}{dt} \frac{1}{2} \int_{\Omega} |w(t)|^2 dx + \int_{\Omega} |\nabla w|^2 dx = \int_S w \partial_\nu w d\sigma + \text{l.o.t.}$$

Write the boundary terms as

$$(8.11) \quad \int_S w \partial_\nu w d\sigma = \int_S w(-N_{\tilde{\Omega}})w d\sigma + \int_S w \text{Op}S^m u_{\text{ex}} d\sigma .$$

If U is the solution of $LU = 0$ in $N_{\tilde{\Omega}}$, $U = 0$ for $t \ll 0$, and $U = w$ on $\partial\Omega$, then

$$\begin{aligned} \int_0^t \int_S w(-N_{\tilde{\Omega}})w d\sigma &\leq \frac{c}{2} \|U(t)\|_{L^2(\tilde{\Omega})}^2 + c \int_0^t \|\nabla U\|_{L^2(\tilde{\Omega})}^2 dt \\ &\leq c \int_0^t \int_S |w|^2 d\sigma dt . \end{aligned}$$

This, along with the lower order terms can be absorbed by a combination of the two terms on the left of (8.10) by an argument we have now used several times.

For the final boundary term in (8.11), observe that u_{ex} extended by zero in $N_{\tilde{\Omega}}$ belongs to $H_{\text{Heat}}^{1+m}([-\infty, T] \times \tilde{\Omega})$ with norm bounded by the $H^m(\Omega)$ norm of u_0 . For $m = 0$, a standard trace inequality shows that $u_{\text{ex}}|_{[0, T] \times S}$ belongs to $H_{\text{Heat}}^{1/2+m}([-\infty, T] \times S)$ with norm bounded by the H^m norm of u_0 . The same result is true for $m = -1$, though in that case H_{Heat}^{1+m} does not have traces. The Sobolev regularity must be combined with an expression for the second normal derivative $\partial_n^2 u_{\text{ex}}$ in terms of tangential and lower order derivatives which is furnished by the equation $Lu = 0$. The steps, in normal coordinates, are as follows. With $I = \{0 < x_n < \eta\}$ so $\mathbb{R} \times \mathbb{R}^{n-1} \times I$ represents a part of $\mathbb{R} \times \tilde{\Omega}$

$$\begin{aligned} u_{\text{ex}} &\in L^2(I : H_{\text{Heat}}^0([-\infty, T] \times \mathbb{R}_{x'}^{n-1})) \\ \partial^2 u_{\text{ex}} / \partial x_n^2 &\in L^2(I : H_{\text{Heat}}^{-2}([-\infty, T] \times \mathbb{R}_{x'}^{n-1})) \\ \partial u_{\text{ex}} / \partial x_n &\in L^2(I : H_{\text{Heat}}^{-1}([-\infty, T] \times \mathbb{R}_{x'}^{n-1})) \\ u_{\text{ex}} &\in C(\bar{I} : H_{\text{Heat}}^{-1/2}([-\infty, T] \times \mathbb{R}_{x'}^{n-1})) . \end{aligned}$$

The integral of the last term of (8.11) is estimated as follows,

$$\begin{aligned} &\int_0^t \int_S w \text{Op}S^m u_{\text{ex}} d\sigma \\ &\leq c \|w\|_{L^2([0, t]; H^{1/2}(S))} \|\text{Op}S^m u_{\text{ex}}\|_{L^2([0, t]; H^{-1/2}(S))} . \end{aligned}$$

Now,

$$\|\text{Op}S^m u_{\text{ex}}\|_{L^2([0, t]; H^{-1/2}(S))} \leq \|\text{Op}S^m u_{\text{ex}}\|_{H_{\text{Heat}}^{-1/2}([-\infty, T] \times S)} .$$

The causality of the operator OpS_{Heat}^m yields

$$\begin{aligned} \|OpS^m u_{\text{ex}}\|_{H_{\text{Heat}}^{-1/2}([-\infty, T] \times S)} &\leq c \|u_{\text{ex}}\|_{H_{\text{Heat}}^{-1/2+m}([-\infty, T] \times S)} \\ &\leq c \|u_{\text{ex}}\|_{H_{\text{Heat}}^m([-\infty, T] \times \tilde{\Omega})} \leq c \|u_0\|_{H^m(\Omega)}. \end{aligned}$$

Summarizing,

$$\int_0^t \int_S w OpS^m u_{\text{ex}} d\sigma \leq c \int_{-\infty}^t \int_{\Omega} |\nabla w|^2 dx dt + c \|u_0\|_{H^m(\Omega)}^2.$$

This estimate together with (8.10), (8.11), and Gronwall's inequality yields the desired result. \square

Acknowledgement. We would like to thank D. Bennequin for his help with the problems of interpolation and E. Dubach for granting permission to cite his interesting computations. We would also like to thank the ex-Centre de Mathématiques Appliquées at the Ecole Normale Supérieure, and the Centres de Mathématiques Pures et Appliquées at the Ecole Polytechnique for their support during the academic years 1985–86 and 1992–1993.

References

1. Beals, R., Fefferman, C. (1974): Spatially inhomogeneous pseudo-differential operators. *Comm. Pure Appl. Math.* **27**, 1–24
2. Dubach, E. (1991): Conditions aux limites artificielles pour les équations de la chaleur et d'advection-diffusion sur une surface fermée – Etude numérique, Prépublication 91-5, Université Paris Nord
3. Engquist, B., Majda, A. (1977): Absorbing boundary conditions for the numerical simulation of waves. *Math. Comp.* **31**, 629–651
4. Givoli, D. (1991): Non-reflecting boundary conditions. *J. Comp. Phys.* **94**, 1–27
5. Hagstrom, T.M. (1991): Asymptotic boundary conditions for dissipative waves: general theory. *Math. Comp.* **56**, 589–606
6. Hagstrom, T.M., Keller, H.B. (1986): The numerical calculation of traveling wave solutions of nonlinear parabolic equations. *SIAM J. Sci. Stat. Comp.* **7**, 978–988
7. Halpern, L. (1986): Artificial boundary conditions for linear advection diffusion equations. *Math. Comp.* **174**, 425–438
8. Halpern, L. (1991): Artificial boundary conditions for incompletely parabolic perturbations of hyperbolic systems. *SIAM J. Math. Anal.* **22**, 1256–1283
9. Hicks, N. (1963): *Notes on Differential Geometry*. Van Nostrand, Princeton NJ
10. Joly, P. (1989): Pseudotransparent boundary conditions for the diffusion equation. Part I, *Math. Meth. Appl. Sci.* **11**, 725–758
11. Lee, J., Uhlmann, G. (1989): Determining anisotropic real analytic conductivities by boundary measurements. *Comm. Pure Appl. Math.* **42**, 1097–1112
12. Lindmann, E.L. (1985): Free-space boundary conditions for the time dependent wave equation. *J. Comp. Phys.* **18**, 16–78
13. Lohéac, J.-P. (1991): An artificial boundary condition for an advection-diffusion equation. *Math. Meth. Appl. Sci.* **14**, 155–175
14. MacDonald, I.G. (1979): *Symmetric Functions and Hall Polynomials*. Oxford, Clarendon Press
15. Nirenberg, L. (1976): *Lectures on Linear Partial Differential Equations*. CBMS Reg. Conf. no. 17, AMS, Providence RI
16. Taylor, M. (1981): *Pseudodifferential Operators*. Princeton University Press, Princeton NJ