

ON THE STUDY OF 1PI ALGORITHMS FOR A GENERAL CLASS OF CURVES

M. KAPRALOV AND A. KATSEVICH

ABSTRACT. We extend an efficient cone beam transform inversion formula, developed earlier by the authors for smooth curves with positive torsion, to a more general class of helix-like curves. These curves are allowed to have negative torsion, and they can be non-smooth at isolated points. The notions of turns and PI segments are extended. The new class is defined by several geometric conditions which impose a tradeoff between the length of critical PI lines (which reflect how severely the positivity of torsion is violated) and the extent to which the curve bends between neighboring turns. The main property of curves from this class is that critical PI lines are allowed to be arbitrarily close to the set U where reconstruction is possible, but are not allowed to intersect it. Some of the conditions that define the class turn out to be common for many known trajectories, so we investigate separately the properties of the Crofton symbol of PI segments of curves that satisfy these conditions. The results of the investigation are then used to develop an efficient filtered backprojection algorithm. Numerical experiments conducted with a clock phantom demonstrate good image quality.

1. INTRODUCTION

Inversion of the cone beam transform arises in pure mathematics (as a problem of integral geometry) and in numerous applications, such as computed tomography (CT). The general statement of the problem is to recover a function $f(x)$, $x \in \mathbb{R}^3$ from values of integrals of f along lines that intersect a curve C . In applications C is usually called a source trajectory. The ever-increasing needs of medical imaging require the development of inversion algorithms for more and more general source trajectories.

Three major groups of theoretically exact reconstruction algorithms have emerged in the past several years: filtered backprojection (FBP) algorithms, slow-FBP algorithms, and backprojection filtration (BPF) algorithms. Slow-FBP and BPF algorithms are quite flexible, allow some transverse data truncation, and can be used for virtually any complete source trajectory [PN05, ZPXW05, SZP05, YZYW05b, YZYW05a, ZLNC04]. FBP algorithms are less flexible, but they are by far the fastest and have been developed for a range of source trajectories. They include constant pitch helix [Kat02, Kat04b, Kat04c, Kat06], dynamic pitch helix [KBH04, KK06], circle-and-line [Kat04a], circle-and-arc [Kat05, CZLN06], and saddle [YLKK06]. Significant progress has also been achieved in the development of quasi-exact algorithms [BKP05, KBK06]. A very nice algorithm has recently been proposed in [PNC05]. The algorithm applies to almost any complete source trajectory and is of the FBP type. However, it sometimes leads to excessive detector requirements.

The work of both authors was supported in part by NSF grant DMS-0505494

M. Kapralov is with the Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305-4042, USA

E-mail address: kapralov@stanford.edu

A. Katsevich is with the Department of Mathematics, University of Central Florida, Orlando, FL 32816-1364, USA

E-mail address: akatsevi@pegasus.cc.ucf.edu.

Recently the algorithm of [Kat04b] has been extended to a general class of smooth curves in [KK07]. It was proved that for any curve C from the class and any point x in the image domain almost all planes passing through x can have only one or three intersection points (IPs) with the PI-segment C_{PI} of x . The main condition characterizing this class is the positivity of torsion. However, since torsion depends on the derivatives of C up to the third order, the condition can be easily violated in practice. Hence it is desirable to find geometric conditions that relax the requirement of positive torsion. This is what we do in the present paper. There are two major new conditions. The first one is based on the notion of a critical chord. A chord H of C is critical if H and vectors tangent to C at both endpoints of H lie in a plane. We distinguish between local and global critical chords. Similarly to [KK07], global critical chords determine generalized turns of C . Local critical chords arise when the torsion of C changes sign. The stronger the violation of the positive torsion condition, the longer the local critical chords can be. In this paper local critical chords are allowed to come arbitrarily close to the set U where reconstruction is possible, but they are not allowed to intersect it. Since local critical chords are related to the torsion of C , our condition is indeed a relaxation of the one in [KK07]. The second new condition controls how the curve bends between neighboring turns. The new class of curves is quite general. In particular, planes through points x in the image domain can have any odd number of IPs with their PI segments (and not just three as in [KK07]). Moreover, these curves are allowed to have discontinuous first derivative at isolated points.

It is also useful to compare the algorithm developed here with the one obtained earlier in [KK06]. First of all, the one in [KK06] applies only to variable pitch helices. Even though both of them apply to helices with non-positive torsion, the new algorithm can handle violations which are significantly more severe than those in [KK06]. For example, the helices admissible in [KK06] always have convex projections inside the field of view, whereas this need not be the case for the new class of curves.

Some of the conditions that define the new class turn out to be common for many known trajectories, so we investigate separately the properties of the Crofton symbol of PI segments of curves that satisfy these conditions. On one hand, this study leads to a method of assigning weights to IPs of planes through x with $C_{PI}(x)$ (cf. the general scheme of [Kat03]). These weights are then used to develop a reconstruction algorithm for the new class. On the other hand, this study is interesting in its own right. In particular, it can be used for developing reconstruction algorithms for classes of curves far different from those described in this paper.

Another important aspect is that the proposed algorithm reduces to that of [Kat04b] when the curve is the standard helix. This is an interesting observation since the definition of weights given here depends on a function $q(\cdot)$ defined on a subset of planes through x . In particular, some filtering planes through x are obtained by solving the equation $q(\Pi) = 0$. Surprisingly, although $q(\cdot)$ depends on the integral of $y(s)$, it turns out that in the case of the standard helix $q(\Pi) = 0$ if and only if Π is the filtering plane through x as defined in [Kat04b], i.e. when Π intersects C at $y(s_0), y(s_1), y(s_2) \in C_{PI}(x)$ with $s_1 = (s_0 + s_2)/2$.

The paper is organized as follows. In section 2 we introduce the new class of curves, define PI lines and maximal PI lines as well as discuss some of their properties. One of the properties turns out to be common for many known trajectories for which 1PI algorithms can be used, so we study the properties of such PI segments in section 3. In section 4 we propose a scheme for assigning weights to IPs. Since the scheme depends on a choice of the function $q(\cdot)$, we consider an example of such function in subsection 4.3 and discuss its properties. The reconstruction algorithm is derived in section 5. Numerical experiments are presented in 6. The main ideas and results of the paper are summarized in section 7. Some proofs and additional remarks are collected in the appendices. For convenience of the reader a list of main symbols is given in appendix D.

2. STRUCTURE OF THE CURVE C

In this section we introduce a class of curves for which we can define the notion of PI-lines similarly to [KK07]. Let C be a continuous piecewise- C^∞ curve with finitely many points of non-smoothness and with no linear segments. We use a non-standard parametrization of C , which is very convenient for our purposes. We describe C by two functions: $y(s)$ and $\tau(s)$:

$$(2.1) \quad I := \cup_{k=1}^K [a_{k-1}, a_k] \ni s \rightarrow y(s) \in \mathbb{R}^3, \quad \tau(s) : I \rightarrow \mathbb{R}^3,$$

with the following properties

- (A1) Each interval $[a_{k-1}, a_k]$ corresponds to either a section where C is C^∞ or to a single point of C where the tangent direction to C is not smooth;
- (A2) If $[a_{k-1}, a_k]$ corresponds to a section where C is C^∞ , then $y(s) \in C^\infty([a_{k-1}, a_k])$ and $\dot{y}(s) \neq 0$, $\tau(s) = \dot{y}(s)$, $s \in [a_{k-1}, a_k]$;
- (A3) If $[a_{k-1}, a_k]$ corresponds to a single point of C , then $y(s) \equiv \text{const}$ on $[a_{k-1}, a_k]$, and

$$(2.2) \quad \tau(s) = \frac{a_k - s}{a_k - a_{k-1}} \tau(a_{k-1}) + \frac{s - a_{k-1}}{a_k - a_{k-1}} \tau(a_k), \quad s \in [a_{k-1}, a_k].$$

Here and below the dot above a variable denotes differentiation with respect to s . Here is some additional notation that is used in the paper. The superscript ‘ $-$ ’ denotes the limit from the left, and ‘ $+$ ’ - the limit from the right. The mixed product of three vectors is denoted $[e_1, e_2, e_3] := e_1 \cdot (e_2 \times e_3)$. As is easily seen, (2.2) describes the vector, which rotates from $\dot{y}(a_{k-1}^-) = \tau(a_{k-1})$ to $\dot{y}(a_k^+) = \tau(a_k^+)$.

A plane is tangent to C at a point $s \in I$ if it contains $y(s)$ and is parallel to $\tau(s)$. Analogously, a line is tangent to C at a point $s \in I$ if it contains $y(s)$ and is parallel to $\tau(s)$. An osculating plane to C at $s \in I$ is the plane $\Pi_{osc}(s)$, which contains $y(s)$ and is parallel to $\tau(s)$ and $\dot{\tau}(s)$. These definitions coincide with the usual ones whenever C is smooth, but they can be applied also at points where C is not smooth.

Given any $s_0, s_1 \in I$, $H(s_0, s_1)$ denotes the line segment with endpoints $y(s_0), y(s_1) \in C$. In view of the non-standard parametrization of C , many different pairs s_0, s_1 may correspond to the same chord. Generally, when we write $H(\cdot, \cdot)$, we tacitly retain the knowledge of which values of the parameter generated that line segment. The only exception to the rule is when we write $H(s_b, s_t)$. In this case $[s_b, s_t]$ is the *shortest* parametric interval that generates the same chord. The same convention applies to curve segments $C(s_0, s_1)$ and $C(s_b, s_t)$.

Following [KK07], introduce the function

$$Q(s_0, s_1) := [y(s_0) - y(s_1), \tau(s_1), \tau(s_0)].$$

As in [KK07], chords $H(s_0, s_1)$ such that $Q(s_0, s_1) = 0$ are called *critical*. Let $H(s_0, s_1)$ be a critical chord. Two cases are possible. If

$$(2.3) \quad \frac{\tau(s_0) \times (y(s_1) - y(s_0))}{|\tau(s_0) \times (y(s_1) - y(s_0))|} = - \frac{\tau(s_1) \times (y(s_1) - y(s_0))}{|\tau(s_1) \times (y(s_1) - y(s_0))|},$$

we call such zeros of Q (and the corresponding critical chords) *local*. If

$$(2.4) \quad \frac{\tau(s_0) \times (y(s_1) - y(s_0))}{|\tau(s_0) \times (y(s_1) - y(s_0))|} = \frac{\tau(s_1) \times (y(s_1) - y(s_0))}{|\tau(s_1) \times (y(s_1) - y(s_0))|},$$

we call such zeros of Q (and the corresponding critical chords) *global*. Consider the plane determined by a critical chord $H(s_0, s_1)$. This plane is tangent to C at s_0 and s_1 . If $H(s_0, s_1)$ is global, then $\tau(s_0)$ and $\tau(s_1)$ point into the same half-plane relative to $H(s_0, s_1)$. If $H(s_0, s_1)$ is local, then $\tau(s_0)$ and $\tau(s_1)$ point into the opposite half-planes relative to $H(s_0, s_1)$.

In order to provide intuition behind the two definitions, it is useful to consider the case of a constant radius/variable pitch helix with torsion that changes sign. Maximal PI lines $H(s, s + 2\pi)$ are global critical chords. They are vertical and define turns of the helix. However, there are other local critical chords $H(s_0, s_1)$ with $s_1 < s_0 + 2\pi$ (see, e.g. [KK06]). These chords generally describe local perturbations of the curve due to, e.g., changes in the sign of torsion.

Remark 2.1. See proposition 2.5, where an important property of global critical chords is established.

For $s \in I$ we denote the first $s' > s$ (resp. $s' < s$) such that $H(s, s')$ (resp. $H(s', s)$) is a global critical chord by $q_{max}(s)$ (resp. $q_{min}(s)$). If no such s' exists, we define $q_{max}(s) := a_K$ (resp. $q_{min}(s) := a_0$). As in [KK07], denote

$$(2.5) \quad N_{min}(s) := y(q_{min}(s)) - y(s), \quad N_{max}(s) := y(q_{max}(s)) - y(s).$$

This allows us to introduce the following definitions:

Definition 2.2. A chord $H(s_0, s_1)$ is a PI line if $s_1 < q_{max}(s_0)$.

Definition 2.3. A chord $H(s_0, s_1)$ is a maximal PI line if $s_1 = q_{max}(s_0)$.

Notations (2.5) and Definitions 2.2, 2.3 resemble those for a conventional helix. If $[s, q_{max}(s)]$ is a turn of C , then a PI line of C is a chord whose endpoints are separated by less than one turn of the curve. For a conventional helix, all ‘‘maximal’’ PI lines $H(s - 2\pi, s)$ and $H(s, s + 2\pi)$ are parallel to the rotation axis. In other words, all the turns have the same orientation (as determined by the line through the endpoints of the turn). This direction is commonly referred to as ‘‘vertical’’. In our case the turns of C can be different, so at each point s we introduce two directions. One, $N_{min}(s)$, points down. The other, $N_{max}(s)$, points up. In particular, each turn of C has its own vertical direction.

If $x \in H(s_b, s_t)$ and $H(s_b, s_t)$ is a PI line in the sense of Definition 2.2, we say that $H(s_b, s_t)$ is the PI line of x . As usual, we denote the PI-parametric interval $[s_b, s_t]$ of x by $I_{PI}(x)$, the PI-segment of C by $C_{PI}(x)$, and the PI line $H(s_b, s_t)$ of x by $L_{PI}(x)$.

We impose the following conditions on the curve C :

- P1. For each $s \in I$, $s \neq a_0, a_K$, all osculating planes $\Pi_{osc}(t)$ satisfy: $[N_{max}(s), \tau(t), \dot{\tau}(t)] > 0$ if $t \in [s, q_{max}(s)]$ and $[N_{min}(s), \tau(t), \dot{\tau}(t)] < 0$ if $t \in [q_{min}(s), s]$;
- P2. For each $s \in I$, $s \neq a_0$, and $t \in [s, q_{max}(s)]$, the tangent line at t does not intersect $C(s, q_{max}(s))$ and the chord $H(s, q_{max}(s))$;
- P3. For each $s \in I$, $s \neq a_0$, the curve segment $C(s, q_{max}(s))$ has no self-intersections.

Conditions P1–P3 are quite natural. For example, P1 says that osculating planes are never vertical. Loosely speaking, P2 implies that, similarly to a helix, C wraps around its ‘interior’.

Proposition 2.4. One has:

- (B1) $q_{max}(s)$ is continuous on $[a_0, a_K)$. $q_{max}(s)$ is differentiable at s if $\dot{\tau}$ is continuous both at s and $q_{max}(s)$. If $q_{max}(a_0) < a_K$, there exists $s_{max}^* \in (a_0, a_K)$ such that $q_{max}(s)$ is strictly increasing and $q_{max}(s) < a_K$ on $[a_0, s_{max}^*)$, and $q_{max}(s) \equiv a_K$ on $[s_{max}^*, a_K]$.
- (B2) Properties of $q_{min}(s)$ are completely analogous. $q_{min}(s)$ is continuous on $(a_0, a_K]$. $q_{min}(s)$ is differentiable at s if $\dot{\tau}$ is continuous both at s and $q_{min}(s)$. If $q_{min}(a_K) > a_0$, there exists $s_{min}^* \in (a_0, a_K)$ such that $q_{min}(s)$ is strictly increasing and $q_{min}(s) > a_0$ on $(s_{min}^*, a_K]$, and $q_{min}(s) \equiv a_0$ on (a_0, s_{min}^*) .
- (B3) $q_{min}(q_{max}(s)) = s$ if $q_{max}(s) < a_K$, and $q_{max}(q_{min}(s)) = s$ if $q_{min}(s) > a_0$;
- (B4) For each $s \in I$ there is no maximal PI line strictly inside $(s, q_{max}(s))$.

Proof. We start with q_{max} . Pick $s_0 \in (a_{k-1}, a_k)$ and assume $q_{max}(s_0) < a_K$. Recall that a_K is the right end-point of I . Similarly to [KK07], consider the equation

$$(2.6) \quad [y(q(s)) - y(s), \tau(q(s)), \tau(s)] = 0, \quad q(s_0) = q_{max}(s_0).$$

Formally differentiating (2.6) with respect to s gives

$$(2.7) \quad q'(s) = -\frac{[y(q(s)) - y(s), \tau(q(s)), \dot{\tau}(s)]}{[y(q(s)) - y(s), \dot{\tau}(q(s)), \tau(s)]} \text{ if } q(s) < a_K.$$

Since $q(s_0) = q_{max}(s_0)$ and $H(s, q(s))$ is a critical chord,

$$N_{max}(s_0) \times \tau(q_{max}(s_0)) = C_1 N_{max}(s_0) \times \tau(s_0),$$

where $C_1 \neq 0$ is some constant. Hence, upon using P1,

$$(2.8) \quad q'(s_0) = \frac{C_1^2 [N_{max}(s_0), \tau(s_0), \dot{\tau}(s_0)]}{[N_{max}(s_0), \tau(q_{max}(s_0)), \dot{\tau}(q_{max}(s_0))]} > 0.$$

By the implicit function theorem there is a unique solution to (2.6) for s close to s_0 . Applying now the same argument as in [KK07], we get that locally $q_{max}(s)$ can be found by (a) finding $q_{max}(s)$ at one point, say s_0 , and (b) extending it to a neighborhood of s_0 by solving (2.6). Hence (2.8) holds not only at s_0 , but also at all s where the extension works, so $q_{max}(s)$ is monotone there. Note that the extension works as long as all the functions involved are smooth.

Consider now two intervals $[a_{k-1}, a_k]$ and $[a_k, a_{k+1}]$, and suppose $\dot{\tau}(s)$ is discontinuous across a_k . Obviously, we do not have to look at discontinuities of higher order derivatives. Find $q_{max}(a_k)$ directly from the definition. Since $y(s)$ and $\tau(s)$ are well-behaved on either side of a_k , we can extend $q_{max}(s)$ from a_k to both intervals (separately) as solutions to (2.6). Using the same argument as in [KK07], it is easy to see that this procedure gives $q_{max}(s)$. Thus, $q_{max}(s)$ is continuous across the boundaries of the intervals. In a similar fashion, $q_{max}(s)$ is continuous across $s = s_0$ even if $\dot{\tau}(q)$ is discontinuous at $q = q_{max}(s_0)$.

The properties of the function $q_{min}(s)$ can be easily obtained from those of $q_{max}(s)$. Denote $q = q_{max}(s)$. Since $H(s, q)$ is a global critical chord, then $q_{min}(q) \geq s$. Suppose there exists another point s_1 , $s < s_1 < q$, such that $H(s_1, q)$ is a global critical chord. Then $q_{max}(s_1) \leq q_{max}(s)$ even though $s_1 > s$. This contradicts the monotonicity of q_{max} , i.e. $q_{min}(q_{max}(s)) = s$ whenever $q_{max}(s) < a_K$. All the remaining assertions of the proposition are now obvious. \square

The following result is very important. Its proof is quite technical and given in appendix A.

Proposition 2.5. *Suppose C satisfies conditions P1-P3. Then for all $s \in I$, $s \neq a_K$, the orthogonal projection of $C(s, q_{max}(s))$ onto $N_{max}^\perp(s)$ is a closed convex curve without self-intersections.*

In the case of a helix, the projection of each turn onto a plane perpendicular to the vertical direction is a circle, i.e. a closed convex curve without self-intersections. Proposition 2.5 shows that in our case the projection of each turn of C along its own vertical direction (see the paragraph following Definition 2.3) has the same property.

Let $\Pi_{max}(s)$ be the plane determined by the critical chord $H(s, q_{max}(s))$: it contains the chord and is parallel to $\tau(s)$ and $\tau(q_{max}(s))$. Similarly, $\Pi_{min}(s)$ denotes the plane determined by the critical chord $H(s, q_{min}(s))$. As an immediate corollary to proposition 2.5 we get

Corollary 2.6. *Fix any $s \in I$, $s \neq a_K$. The curve segment $C(s, q_{max}(s))$ stays on one side of $\Pi_{max}(s)$. Similarly, for any $s \in I$, $s \neq a_0$, the curve segment $C(q_{min}(s), s)$ stays on one side of $\Pi_{min}(s)$.*

As in [KK07], it follows that for each $s \in I$ there are two ‘detector planes’ - $DP_+(s)$ and $DP_-(s)$ - on which the upper and lower ‘turns’ of C can be projected with only one asymptote. $DP_+(s)$ is any plane not passing through $y(s)$ and parallel to $\Pi_{max}(s)$. Similarly, $DP_-(s)$ is any plane not passing through $y(s)$ and parallel to $\Pi_{min}(s)$. In principle, the exact location of $DP_+(s)$ and $DP_-(s)$ is irrelevant. For simplicity, we assume here and in what follows that $DP_+(s) \cap DP_-(s)$ is a line that belongs to $\Pi_{osc}(s)$ and $\hat{\tau}(s)$ points towards that line. It is clear that this line is parallel to $\tau(s)$. We shall denote this line by L_0 . The stereographic projection of $C(q_{min}(s), s)$ from $y(s)$ onto $DP_-(s)$ is denoted Γ_- , and the stereographic projection of $C(s, q_{max}(s))$ from $y(s)$ onto $DP_+(s)$ is denoted Γ_+ . The following result is proven analogously to proposition 4.4 of [KK07].

Proposition 2.7. Γ_- and Γ_+ are continuous and piecewise- C^∞ . They are nonsmooth at most at those s' where C is nonsmooth.

Until now we have been discussing the properties of C within one turn, i.e. $C(s, q_{max}(s))$. Next we discuss the relationship between neighboring turns of C , i.e. the curve segments $C(q_{min}(s), s)$ and $C(s, q_{max}(s))$. It is convenient to consider the orthogonal projection of $C(q_{min}(s), q_{max}(s))$ onto the plane $\tau(s)^\perp$, which we denote by \hat{C} (see Fig. 1). By property P1, $q_{min}(s)$ is always located below $\Pi_{osc}(s)$ and $q_{max}(s)$ is located above it. Here ‘above’ and ‘below’ are determined relative to the direction of the normal vector $\tau(s) \times \hat{\tau}(s)$, which is used in property P1. By corollary 2.6, \hat{C} is on the side of the lines given by $N_{max}(s)$ and $N_{min}(s)$ (i.e., $\Pi_{max}(s) \cap \tau^\perp(s)$ and $\Pi_{min}(s) \cap \tau^\perp(s)$) where $\hat{\tau}(s)$ points. Hence only the following two cases are possible:

- (C1) the angle between the vectors $\hat{N}_{max}(s)$ and $\hat{N}_{min}(s)$ is less than π (see Fig. 1, left panel);
- (C2) the angle between the vectors $\hat{N}_{max}(s)$ and $\hat{N}_{min}(s)$ is greater than π (see Fig. 1, right panel).

In order to gain insight into the additional conditions proposed below, it is useful to consider the class of curves with positive torsion from [KK07]. This is a fairly general class of curves, but their properties with respect to exact and efficient image reconstruction are fairly simple. Here we propose a much wider class, which is a generalization of the one in [KK07] and allows for a much more complicated curves.

For the curves from [KK07] we have: $C(s, q_{max}(s))$ is above $\Pi_{osc}(s)$, $C(q_{min}(s), s)$ is below $\Pi_{osc}(s)$, and $\Pi_{osc}(t)$, $t \in (q_{min}(s), q_{max}(s)) \setminus \{s\}$ never passes through $y(s)$. Hence, if case (C1) happens (see Fig. 1, left panel), $\hat{C}(s, q_{max}(s))$ never intersects the line determined by $\hat{N}_{min}(s)$. However, if the torsion is allowed to be negative, the curve may go below $\Pi_{osc}(s)$ and even intersect the line determined by $\hat{N}_{min}(s)$. It is also important to note that \hat{C} is tangent to $\hat{\tau}(s)$ with a cusp, so it cannot intersect $\hat{N}_{min}(s)$ close to s . This leads us to the following condition, which is a relaxation of the conditions imposed on the curve in [KK07]:

- P4(1) If case (C1) happens, $C(s, q_{max}(s))$ does not intersect $\Pi_{min}(s)$, and $C(q_{min}(s), s)$ does not intersect $\Pi_{max}(s)$.

Define the lines $L_{min}^+ := DP_+(s) \cap \Pi_{min}(s)$ on $DP_+(s)$ and $L_{max}^- := DP_-(s) \cap \Pi_{max}(s)$ on $DP_-(s)$. If case (C1) happens, L_{min}^+ is below L_0 on $DP_+(s)$, and L_{max}^- is above L_0 on $DP_-(s)$ by P1 (see Fig. 1, left panel). Property P4(1) asserts that Γ_+ stays above L_{min}^+ on $DP_+(s)$, and Γ_- stays below L_{max}^- on $DP_-(s)$. This is another way to see that we significantly relaxed conditions on the curve compared with [KK07]. Indeed, if the curve does not bend too much, then L_{min}^+ is far below L_0 on $DP_+(s)$, and L_{max}^- is far above L_0 on $DP_-(s)$. At the same time, if C satisfies the conditions in [KK07], Γ_+ is above L_0 , and Γ_- is below L_0 . Hence a significant perturbation of the curve is required to bring Γ_+ below L_{min}^+ or to bring Γ_- above L_{max}^- .

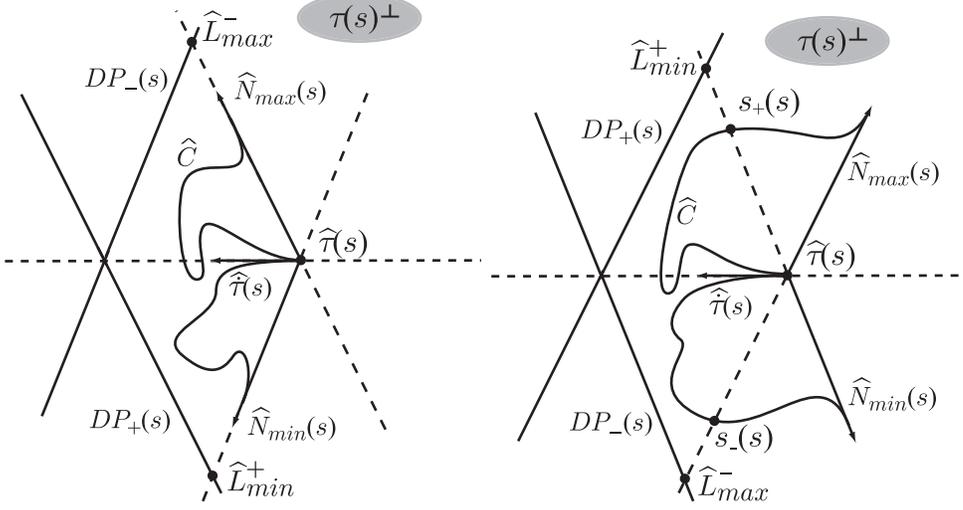


FIGURE 1. Relative position of $N_{min}(s)$ and $N_{max}(s)$. Case (C1) - left panel, case (C2) - right panel.

We now consider the behavior of curves from [KK07] when case (C2) happens (see Fig. 1, right panel). One sees that $\hat{C}(s, q_{max}(s))$ necessarily intersects the line determined by $\hat{N}_{min}(s)$ at least once, which leads us to the second condition:

P4(2). Denote the first intersection point (IP) of $C(s, q_{max}(s))$ with $\Pi_{min}(s)$ by $q_+(s)$, and the first IP of $C(q_{min}(s), s)$ with $\Pi_{max}(s)$ by $q_-(s)$. If case (C2) happens, then $q_+(s)$ and $q_-(s)$ are the only IPs. Also, $y(s) \notin \Pi_{osc}(t)$ for all $t \in (q_{min}(s), q_-(s)) \cup (q_+(s), q_{max}(s))$.

If case (C1) happens, then for simplicity we define $q_+(s) := q_{max}(s)$ and $q_-(s) = q_{min}(s)$. It is important to note here that, in a sense, condition P4(2) complements condition P1. P1 implies that $\Pi_{osc}(q_{max}(s))$ does not contain $y(s)$. If case (C2) occurs, condition P4(2) requires that all $\Pi_{osc}(q)$, $q \in [q_+(s), q_{max}(s)]$, do not contain $y(s)$. Of course, a similar remark regarding $q_{min}(s)$ is true as well.

Following [KK07], define

$$(2.9) \quad \Phi(t, s) := [y(t) - y(s), \tau(t), \dot{\tau}(t)].$$

It is easy to see that $\Phi(t, s)$ is proportional to the directed distance from $y(s)$ to $\Pi_{osc}(t)$. Property P4(2) implies that $\Phi(t, s) > 0$ for $t \in (q_+(s), q_{max}(s))$, and $\Phi(t, s) < 0$ for $t \in (q_{min}(s), q_-(s))$. One can show in the same way as in [KK07] (see the proof of proposition 4.4 in [KK07]) that if C is smooth at t , then up to a positive factor $\Phi(t, s)$ gives the curvature of Γ_+ on $DP_+(s)$ and of Γ_- on $DP_-(s)$ at $y(t)$. Thus, by property P4(2), the sections of Γ_+ and Γ_- located close to $q_{max}(s)$ and $q_{min}(s)$, respectively, are convex. Because of our assumptions, the convexity is maintained even at points of nonsmoothness.

Let us compare condition P4(2) with conditions in [KK07]. If case (C2) holds, L_{min}^+ is above L_0 on $DP_+(s)$, and L_{max}^- is below L_0 on $DP_-(s)$ by P1. If C satisfies the conditions in [KK07], the main of them being positive torsion, then Γ_+ is above L_0 on $DP_+(s)$, Γ_- is below L_0 on $DP_-(s)$, and both curves are convex. Consequently, L_{min}^+ and L_{max}^- intersect Γ_\pm only once, so property P4(2) holds. In contrast with [KK07], here we require that only the portion of Γ_+ above L_{min}^+ and the portion of Γ_- below L_{max}^- are convex. If the curve does not bend too much between neighboring turns, both L_{min}^+ and L_{max}^- are far away from L_0 and those portions are quite small in

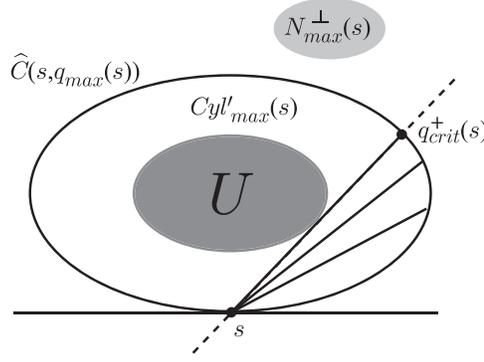


FIGURE 2. Illustration of $Cyl_{max}(s)$ and $Cyl'_{max}(s)$ on $N_{max}^{\perp}(s)$

the sense of parameter range: $q_+(s)$ is close to $q_{max}(s)$, and $q_-(s)$ is close to $q_{min}(s)$. By condition P1, $\Pi_{osc}(q_{max}(s))$ and $\Pi_{osc}(q_{min}(s))$ do not contain $y(s)$. It is therefore natural to expect that the planes $\Pi_{osc}(q)$ do not contain $y(s)$ for q close to $q_{max}(s) \cup q_{min}(s)$. Similarly, if C does not bend too much between neighboring turns, the lines determined by $\tilde{N}_{max}(s)$ and $\tilde{N}_{min}(s)$ (cf. Figure 1) are close to each other, and only a significant perturbation of the curve can cause these lines to have more than one IP with \hat{C} . So our requirements again significantly relax those in [KK07].

As is seen from the discussion, there are two major ways to make a curve from [KK07] more complicated. One is to violate the positivity of torsion, and the other is to allow the curve to bend very rapidly between neighboring turns. Condition P4 combines the two phenomena into a single geometrical requirement. Thus, we do not need to impose two separate conditions: one on torsion, and the other on the behavior of neighboring turns. This means, for example, that if the torsion condition is strongly violated, then, as compensation, the curve should bend very little between neighboring turns so that overall P4 is still satisfied.

It remains to define a set $U \subset \mathbb{R}^3$ where reconstruction is possible with an efficient algorithm. Since our algorithm will be PI line based, it is natural to require that PI lines exist and are unique for all $x \in U$. In what follows we define U in a way that ensures local uniqueness of PI lines. Existence can be achieved in a standard fashion by restricting U to the region covered by PI lines of C (see section 6).

As in [KK07], let $Cyl_{max}(s)$ be the infinite cylinder with base $\hat{C}(s, q_{max}(s))$ and axis $N_{max}(s)$. We need to shrink $Cyl_{max}(s)$ to get rid of local critical chords. For each $s \in I$ consider local critical chords $H(s, s')$, $s < s' < q_{max}(s)$, and denote the maximal s' by $q_{crit}^+(s)$. From property P1, $q_{crit}^+(s) < q_{max}(s)$. Finally, let $Cyl'_{max}(s)$ be the part of the cylinder $Cyl_{max}(s)$ that is cut off from it by the plane through $y(q_{crit}^+(s))$ and $y(s)$ and containing $N_{max}(s)$ (see Fig. 2). In a completely similar fashion we define $q_{crit}^-(s)$ and construct the cylinders $Cyl'_{min}(s)$.

We now proceed in exactly the same fashion as in [KK07], i.e. choose an interval $I' \subset I$ and define U' as

$$(2.10) \quad U' := \bigcap_{s \in I'} (Cyl'_{max}(s) \cap Cyl'_{min}(s)).$$

If I' is sufficiently short, then $U' \neq \emptyset$ (see the discussion after the proof of Proposition 3.2 in [KK07]). Definition (2.10) ensures that there is no more than one PI line through $x \in U'$ with at least one endpoint in I' . This can be shown in the same way as in section 3 of [KK07]. We describe the construction here for convenience of the reader. Due to (2.10), all $x \in U'$ are projected inside the closed convex curve $\hat{C}(s, q_{max}(s))$. Hence, one can find $s < q(s) < q_{max}(s)$ and $0 < \lambda(s) < 1$

such that for $x(s) := y(s) + \lambda(s)(y(q(s)) - y(s))$ one has

$$x(s) - x \parallel N_{max}(s).$$

Define $\varepsilon(s) := N_{max}(s) \cdot \{(y(s) + \lambda(s)(y(q(s)) - y(s))) - x\}$. It is shown in [KK07] that the condition $Q(s, q(s)) \neq 0$, which is guaranteed for all $s \in I'$ because $x \in U$, implies that $\varepsilon(s) = 0$ cannot have more than one solution in I' . This, in turn, implies that there cannot be more than one PI line through $x \in U'$ with $s_b \in I'$. To prove that there cannot be more than one PI line through $x \in U'$ with $s_t \in I'$ we argue analogously using the directions $N_{min}(s)$, $s \in I'$.

We now define the set U where reconstruction will be shown to be possible with an FBP-type algorithm as the set of $x \in U'$ with $I_{PI}(x) \subset [c, d]$, i.e. we define

$$(2.11) \quad U := U' \cap \{x \in H(s_0, s_1) \mid c \leq s_0 < s_1 \leq d\}$$

and introduce the condition

P5. The set U defined in (2.11) is nonempty.

Note that $U' \neq \emptyset$ does not imply $U \neq \emptyset$. For example, if I' is so short or local critical chords are so long that $d < q_{crit}^+(c)$, then U can be empty. That is why we use U instead of U' in condition P5.

The last assumption we need to make is that the curve C does not oscillate too much.

P6. For each $x \in U$ and $s \in I_{PI}(x)$ the number of planes that contain x and $y(s)$ and are tangent to $C(q_{min}(s), q_{max}(s))$ is uniformly bounded.

We conclude the section by noting the following property of PI lines of curves from the class that we have introduced. For each $x \in U$ with PI line $H(s_b, s_t)$ the line segment joining \hat{x} and $\hat{y}(t)$, $t \in [s_b, s_t]$, on the projection onto $N_{max}^\perp(s_b)$ rotates monotonically and sweeps π radians (as follows from proposition 2.5 and corollary 2.6). In the next section we focus specifically on curve segments $C(s_b, s_t)$ possessing the property that for a chosen $x \in H(s_b, s_t)$ there exists a vector D such that the line segment joining \hat{x} with $\hat{y}(t)$ in the plane D^\perp rotates monotonically and sweeps π radians as t goes from s_b to s_t .

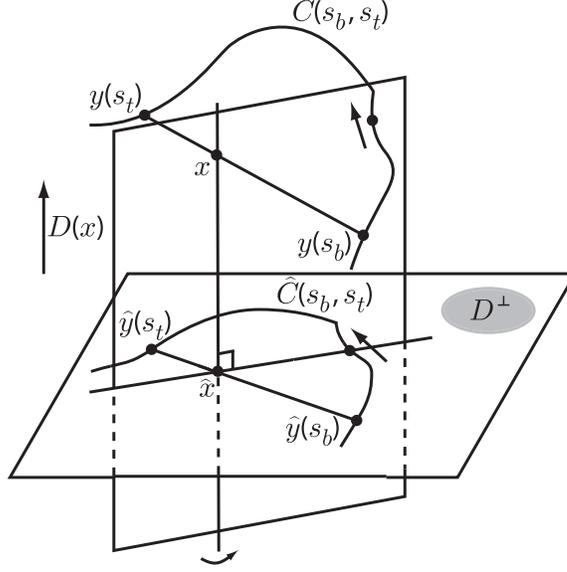
3. PROPERTIES OF 1PI CURVE SEGMENTS

The objective of this section is to define a general class of curves along the lines of the property noted at the end of the previous section and study some of its properties. We start from scratch by ignoring properties P1–P6 (in particular, we allow linear segments). We do this because the condition by itself is quite important and leads to interesting observations, so we felt the need to single it out and conduct an investigation independent of all the other conditions stated in Section 2.

According to the Radon transform theory, it is important to know how many IPs there are between any plane Π through x and $C_{PI}(x)$. This number is known as the Crofton symbol. Here we develop a geometric construction, which allows us to capture the entire dynamics of the Crofton symbol. Clearly, every plane Π through x can be described by a unit vector. When Π changes, i.e. the corresponding point on the unit sphere moves, the number of IPs in $\Pi \cap C_{PI}(x)$ may change only when Π contains $L_{PI}(x)$ or when Π is tangent to $C_{PI}(x)$. These critical planes correspond to some curves on the unit sphere. We establish that the existence of a PI line puts some restrictions on the shape of those curves. Additionally, the construction allows us to classify (almost all) planes through x as *proper* or *improper*. Such classification is then used in Section 4 for developing a reconstruction algorithm.

Let C be a continuous piecewise C^∞ curve given by (2.1).

Definition 3.1. Fix a chord $H(s_b, s_t)$ and point $x \in H(s_b, s_t)$. The chord $H(s_b, s_t)$ is called a PI line of x if a vector $D(x) \in S^2$ can be chosen so that

FIGURE 3. Projection of the PI line along D

- (D1) $[y(s) - x, \tau(s), D(x)] > 0$ for all $s \in [s_b, s_t]$, and
(D2) any plane through x that is parallel to $D(x)$ and does not contain $H(s_b, s_t)$, intersects $C(s_b, s_t)$ at precisely one point.

This definition is quite natural and matches precisely the property mentioned at the end of last section. Consider a plane which contains x and rotates around the vector $D(x)$ (see Fig. 3). Then $H(s_b, s_t)$ is a PI line of $x \in H(s_b, s_t)$ if the point of intersection of the plane and $C(s_b, s_t)$ moves in one direction from $y(s_b)$ to $y(s_t)$.

Note that this coincides with the standard definitions of PI lines for known trajectories, such as constant- and dynamic-pitch helices (choose $D(x) \equiv (0, 0, 1)$ for all x), circle and line (choose a slightly tilted $(0, 0, 1)$), as well as the general class of curves described in [KK07] (here one can choose $N_{max}(s_b(x))$, where $s_b(x)$ is the lower endpoint of the PI line through x).

Our first assumption on the source trajectory is as follows.

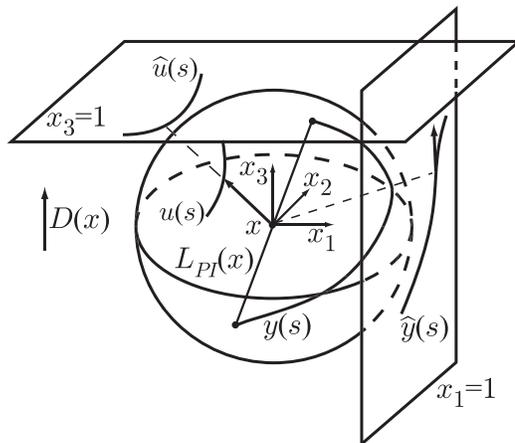
- P1'. There exists an open set $U \subset \mathbb{R}^3$ with the property that all $x \in U$ possess locally unique PI lines and the vector $D(x)$ depends continuously on x .

Proposition 2.5 and Corollary 2.6 imply that curves from the class introduced in section 2 satisfy property P1'.

Let $\Omega_{crit}(x)$ be the set of planes through x that are tangent to $C_{PI}(x)$, or contain $L_{PI}(x)$, or contain a point where $C_{PI}(x)$ is not smooth. As is well-known, $\Omega_{crit}(x)$ is a set of measure zero. Let $\Pi_*(x)$ denote the set of planes through x which are not in $\Omega_{crit}(x)$ and have three or more IPs with $C_{PI}(x)$.

Corollary 3.2. For all $x \in U$ one can choose normals $N(x, \Pi) \in S^2$ of planes $\Pi \in \Pi_*(x)$ so that $N(x, \Pi)$ is continuous and $D(x) \cdot N(x, \Pi) > 0$.

Proof. By Definition 3.1, all planes passing through x , containing $D(x)$, and not containing $L_{PI}(x)$, have only one IP with $C_{PI}(x)$. Hence, $\Pi_*(x)$ corresponds to two sets in S^2 – one in the hemisphere $\{\xi \in S^2 | D(x) \cdot \xi > 0\}$, and the other in the hemisphere $\{\xi \in S^2 | D(x) \cdot \xi < 0\}$.


 FIGURE 4. Construction of the T -curve in the plane $x_3 = 1$

Hence, one can choose the normals $N(x, \Pi)$ for $\Pi \in \Pi_*(x)$ so that $N(x, \Pi)$ is continuous in both x (by the continuity of $D(x)$) and $\Pi \in \Pi_*(x)$. \square

In the remainder of this section a point $x \in U$ is fixed, so the endpoints of $I_{PI}(x)$ are denoted by s_b and s_t . In subsequent derivations we use a number of geometric constructions that have been introduced in [Kat06]. For convenience of the reader they are repeated here.

The T -curve of x is defined in [Kat06] as

$$(3.1) \quad u(s) := \frac{(y(s) - x) \times \tau(s)}{|(y(s) - x) \times \tau(s)|}, \quad s \in I_{PI}(x).$$

Choose the coordinate system (x_1, x_2, x_3) so that x is the origin, $D(x)$ points along the x_3 -axis, and the vector $(y(s_t) - y(s_b)) \times D(x)$ points along the x_1 -axis. Clearly, all curves $y_1(s) := \kappa(s)y(s)$, where $\kappa(s) > 0$ is an arbitrary continuous piecewise C^∞ function, have the same T -curve. Hence, for simplicity, we can assume that C belongs to the unit sphere, i.e. $|y(s)| = 1$. By construction, $C_{PI}(x)$ lies in the halfspace $x_1 \geq 0$. Therefore, using the stereographic projection from the origin, $C_{PI}(x)$ can be associated with a curve in the plane $x_1 = 1$ going from negative infinity to positive infinity along the x_2 -axis. We denote this curve by $\hat{y}(s)$ (see Fig. 4). Definition 3.1 implies that lines tangent to $\hat{y}(s)$ are never vertical. Thus the T -curve is traced by $u(s)$ on the upper hemisphere and, hence, can be associated with a curve $\hat{u}(s)$ in the plane $x_3 = 1$ using stereographic projection. Our construction identifies lines tangent to the curve $\hat{y}(s)$ with directions $u(s) \in S^2$ or with points $\hat{u}(s)$ in the plane $x_3 = 1$ (see Fig. 4). It is easy to see using (A3) and (2.2) that a point where C is not smooth corresponds to a line segment in $\hat{u}(s)$ (see Fig. 5).

Here is a summary of the construction. Given a point $y(s) \in C$, consider a plane through the origin, $y(s)$, and parallel to $\tau(s)$. The same plane is defined by $\hat{y}(s)$ and $\hat{\tau}(s)$. Then $u(s)$ is the normal vector to the plane.

On the other hand, we can start with the curve $u(s) \in S^2$ defined by (3.1). If $u(s)$ is smooth,

$$(3.2) \quad u(s) \times \dot{u}(s) = c(s)[y(s), \tau(s), \dot{\tau}(s)]y(s),$$

where $c(s) > 0$. Hence we can identify lines tangent to the curve $\hat{u}(s)$ with directions $y(s) \in S^2$ or with points $\hat{y}(s)$ in the plane $x_1 = 1$. This is how it works. For each s_0 where $\hat{u}(s)$ is smooth consider the plane passing through the origin and tangent to the curve at $\hat{u}(s_0)$. The unit vector $N = (N_1, N_2, N_3)$, which is perpendicular to the plane and satisfies $N_1 > 0$, coincides with $y(s_0)$.

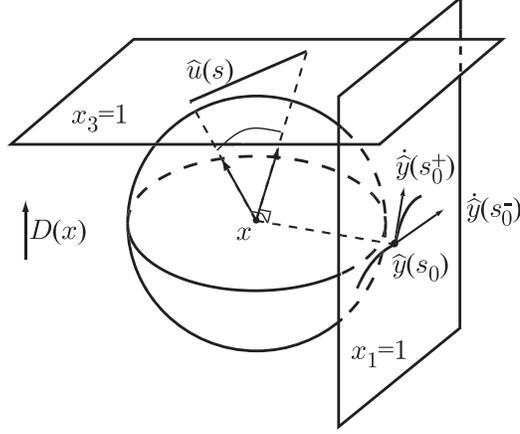


FIGURE 5. A discontinuity of $\hat{\tau}(s)$ corresponds to a linear segment of $\hat{u}(s)$

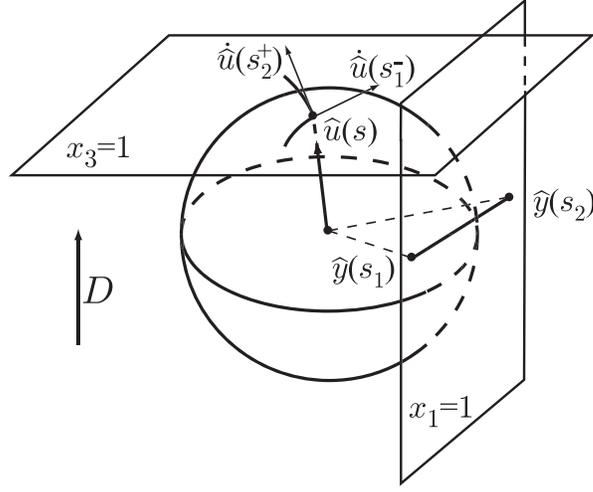


FIGURE 6. A linear segment of the projected curve corresponds to a discontinuity of $\hat{u}(s)$

In other words, the T -curve of $u(s)$ is the original curve $y(s)$. Such a duality is not surprising in view of the involutivity of the Legendre transform.

Suppose $\hat{u}(s)$ is not smooth at some s_0 , but the limiting tangent vectors $\hat{u}(s_0^-)$ and $\hat{u}(s_0^+)$ are parallel. For example, \hat{u} may have a cusp at s_0 . Then the two vectors determine the same plane, and $y(s_0)$ coincides with the normal vector to that plane. Suppose next that $u_0 \in u(s)$ is a corner, i.e. the limiting tangents from the left and from the right at u_0 are not parallel (see Fig. 6). Equation (3.2) implies that the only time this can happen is when $[y(s), \tau(s), \dot{\tau}(s)] \equiv 0$ over some interval, say $[s_1, s_2]$. Equivalently, this means that $x \in \Pi_{osc}(s)$ for all $s \in [s_1, s_2]$, or $y(s), s \in [s_1, s_2]$, is an arc of a great circle in the hemisphere $x_1 > 0$, or $\hat{y}(s), s \in [s_1, s_2]$, is a line segment (see Fig. 6). Hence the missing part of $y(s)$ can be recovered by connecting points $y(s_1)$ and $y(s_2)$, which are determined using $\hat{u}(s_1^-)$ and $\hat{u}(s_2^+)$, by an arc of a great circle.

In summary, $y(s)$ can be found using the following steps:

- (E1) Find $y(s)$ using (3.2) and the constraints $y(s) \in S^2$, $y_1(s) > 0$, at all the points where $u(s)$ is smooth;
- (E2) Fill in all the single-point gaps by continuity; and
- (E3) Fill in the remaining gaps by connecting the corresponding endpoints by an arc of the great circle.

If $\hat{u}(s)$ is nonsmooth at \hat{u}_0 , the two limiting tangents determine two wedges with vertex \hat{u}_0 in the plane $x_3 = 1$. Choose the wedge with an opening less than π and assume that the line tangent to $\hat{u}(s)$ at \hat{u}_0 rotates from the limiting left position to the limiting right position within the smaller wedge. In other words, we augment the set of lines tangent to $\hat{u}(s)$ by lines located within the smaller wedge at every point where $\hat{u}(s)$ is not smooth, thereby extending the conventional definition of a tangent line. It is easy to see that this extended definition achieves two goals: (a) lines tangent to $\hat{u}(s)$ form a continuous family, and (b) using (E1)–(E3) this family is continuously mapped into the curve $\hat{y}(s)$ (see Proposition 3.4 below).

We need two more definitions from [Kat06]. The great circle in S^2 that corresponds to planes through x containing $H(s_b, s_t)$, i.e. $(y(s_t) - y(s_b))^\perp \subset S^2$ is called the *A-curve* of x . In the coordinate system that we chose the *A-curve* corresponds to the *A-line* in the plane $x_3 = 1$, which is parallel to the x_1 -axis. Given $s \in I_{PI}(x)$, the set of planes through x and $y(s)$, i.e. the great circle $(y(s) - x)^\perp \subset S^2$, is called the *B_s-curve*. Hence, the *B_s-curve* corresponds to a line in the plane $x_3 = 1$. We call it a *B_s-line*. It was shown in [Kat06] for the standard helix and will be shown below for more general curves that the *B_s-lines* constitute the set of all tangent lines to the *T-curve* $\hat{u}(s)$. In the same fashion, $\Pi_*(x)$ can also be viewed as a subset of the plane $x_3 = 1$. Property P1 ensures that Π_* is bounded there.

Next we characterize *T-curves* of PI segments of curves that satisfy our assumptions. As before, we choose a coordinate system so that x is the origin and assume $|y(s)| \equiv 1$. Recall that we use the extended definition of a family of lines tangent to $\hat{u}(s)$.

Proposition 3.3. *A curve $T \subset S^2$ is the T-curve for $C_{PI}(x)$ if and only if T is a continuous piecewise C^∞ curve satisfying the following conditions:*

- (F1) *there exists a unit vector $D \in S^2$ such that $u \cdot D > 0$ for all $u \in T$;*
- (F2) *there exists a great circle C_A such that u starts and ends on C_A and is tangent to it at the endpoints;*
- (F3) *the tangent line to the projection of T onto D^\perp rotates counterclockwise and sweeps π radians as the point of tangency moves from one endpoint to the other.*

The proof of the proposition is given in appendix B. The following result describes the main properties of *B_s-curves* (*B_s-lines*). Its proof is also given in appendix B.

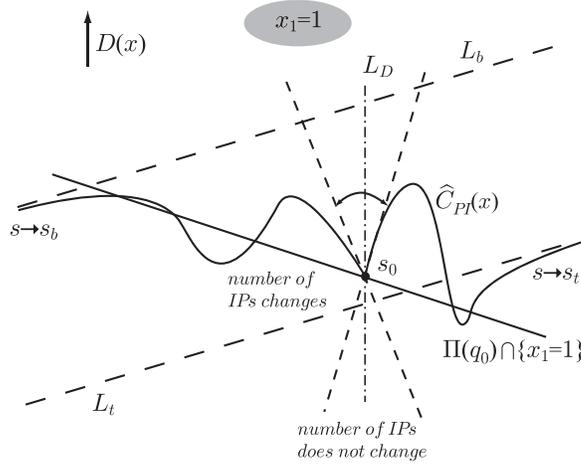
Proposition 3.4. *One has:*

- (G1) *The set of B_s-lines is precisely the set of all tangent lines to the T-curve;*
- (G2) *Lines tangent to the T-curve cover the whole plane $x_3 = 1$; and*
- (G3) *If Π is a plane through x and $\Pi \notin \Omega_{crit}(x)$, then the number of IPs in $\Pi \cap C_{PI}(x)$ is equal to the number of tangents to the T-curve that pass through the point corresponding to Π in the plane $x_3 = 1$.*

Note that proposition 3.4 describes properties of the Crofton symbol of x , which is defined as

$$(3.3) \quad Cr(x, \xi) = \#\{y \in C_{PI}(x) | (y - x) \cdot \xi = 0\}.$$

For example, (G2) is another way of stating that the set $\{\xi \in S^2 | Cr(x, \xi) = 0\}$ is empty, i.e. $C_{PI}(x)$ is complete. (G3) states that the boundary of sets $\{\xi \in S^2 | Cr(x, \xi) \geq 2k + 1\}$ has negative curvature, and exactly $2k + 1$ tangents to the *T-curve* (the latter is the boundary between the regions with different k) intersect at each point $\xi \in S^2$ such that $Cr(x, \xi) = 2k + 1$.

FIGURE 7. Discontinuity of $\hat{\tau}(s)$

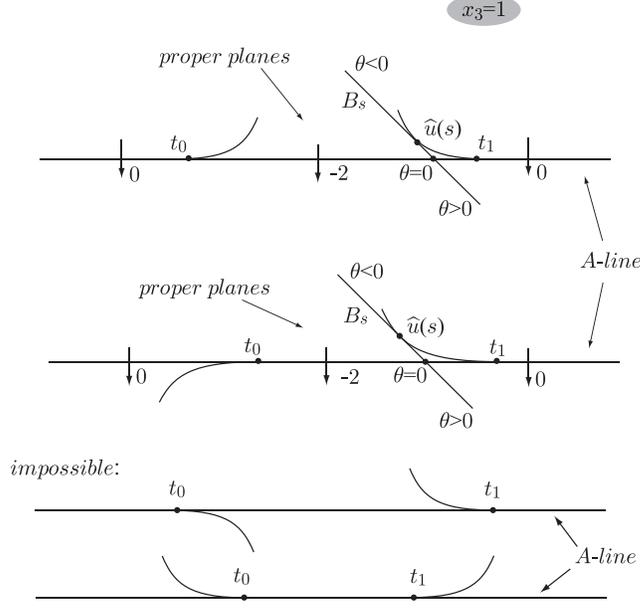
Let $\Pi(q)$ be a smooth one-parametric family of planes through x . We need to study how the number of IPs in $\Pi(q) \cap C_{PI}(x)$ changes with q . It is clear that the number of IPs can change at some $q = q_0$ only when $\Pi(q_0)$ contains $L_{PI}(x)$, or $\Pi(q_0)$ is tangent to $C_{PI}(x)$ at a point where $\dot{y}(s)$ is continuous, or $\Pi(q_0)$ contains a point $y(s)$ where $\dot{y}(s)$ is discontinuous. Identify $\Pi(q)$ with a curve on the unit sphere (via its normal vector) and, consequently, with a curve in the plane $x_3 = 1$. Thus the first event occurs when $\Pi(q)$ intersects the A -line, the second – when $\Pi(q)$ intersects the T -curve, and the third – when $\Pi(q)$ intersects the line spanned by a linear segment of the T -curve (cf. the discussion following Corollary 3.2).

Suppose $y_0 \equiv y(s)$ for all $s \in [a_{k-1}, a_k]$ and $y_0 \in \Pi(q_0)$ for some q_0 . Then the lines through \hat{y}_0 and parallel to $\hat{\tau}(a_{k-1})$ and $\hat{\tau}(a_k)$ form two wedges in the plane $x_1 = 1$ (see Fig. 7). One of the two wedges contains the line L_D through \hat{y}_0 and parallel to $D(x)$. It is clear that the number of IPs in $\Pi(q) \cap C_{PI}(x)$ does not change across q_0 if the line $\Pi(q_0) \cap \{x_1 = 1\}$ belongs to the wedge where L_D is located. Otherwise, the number of IPs changes by two. From the discussion preceding Proposition 3.3, this implies that the number of IPs changes only if the curve intersects the linear segment of the T -curve corresponding to $\hat{y}(s_0)$. This argument implies that the number of IPs may change only if $\Pi(q)$ intersects the A -line or the T -curve (as defined prior to Proposition 3.3). Hence the T -curve and the A -line divide the plane $x_3 = 1$ into regions D_j , $j \geq 1$, such that all planes inside each D_j have the same number of IPs with $C_{PI}(x)$. When a plane moves from one D_j into another across the T -curve, it either gains or loses two IPs through tangency. However, the behavior can be different when the plane moves from one D_j into another across the A -line.

We now consider how the number of IPs changes across the A -line. Introduce a coordinate t along the A -line and denote the coordinates of the two points where the T -curve is tangent to the A -line by t_0 and t_1 (see Fig. 8). In this figure θ is a parameter along B_s , which is used in section 4 below (see the paragraph preceding (4.1)). The two points t_0 and t_1 correspond to the planes containing $L_{PI}(x)$ and tangent to $C_{PI}(x)$ at s_b and s_t respectively. We need to make sure that $t_0 \neq t_1$, i.e. that the two planes are distinct. This leads us to the condition that the PI line of x should not be *critical* (see [KK06, KK07]), i.e.

$$P2'. [y(s_t) - y(s_b), \tau(s_b), \tau(s_t)] \neq 0 \text{ for all } x \in U,$$

where $U \subset \mathbb{R}^3 \setminus C$ is an open set. Again, definition 2.2 and (B4) imply that if the set U defined in (2.11) is non-empty, then it satisfies property P2'.


 FIGURE 8. Number of IPs near the A -line

By proposition 3.4, the number of IPs of a given plane with $C_{PI}(x)$ is equal to the number of tangents to the T -curve passing through the corresponding point in the plane $x_3 = 1$. Hence the change in the number of IPs from one side of the A -line to the other side depends on the location of points t_0 and t_1 , since this is where new lines tangent to the T -curve emerge (see Fig. 8). By proposition 3.3, the T -curve is tangent to the A -line at $\hat{u}(s_b)$ and $\hat{u}(s_t)$ (which correspond to coordinates t_0 and t_1). By the same proposition, the tangent line to $\hat{u}(s)$ sweeps an angle of π radians as s increases from s_b to s_t . Thus the T -curve cannot be tangent to the A -line at any point except for $\hat{u}(s_b)$ and $\hat{u}(s_t)$. Hence we need to consider only the behavior of the T -curve near $\hat{u}(s_b)$ and $\hat{u}(s_t)$. One sees that the number of IPs depends on the side of the A -line on which the T -curve stays after tangency as well as on the direction of the tangent vectors $\dot{\hat{u}}(s_b)$ and $\dot{\hat{u}}(s_t)$.

Examining the possibilities and taking into account property P1' shows that (a) the number of IPs does not change across the A -line if the point of crossing is outside the interval $[t_0, t_1]$, and (b) the number of IPs changes by two when the A -line is crossed inside the interval $[t_0, t_1]$. Two top panels of Fig. 8 show the only two possible configurations of tangency at s_b and s_t (up to a reflection). Two bottom panels of Fig. 8 show impossible configurations. They are impossible because they contradict the fact that the tangent line to the T -curve sweeps an angle of π as s goes from s_b to s_t (the angle would have to be greater in this case). In the two upper diagrams of Fig. 8 the arrows crossing the A -line with numbers near them denote the change in the number of IPs.

Let Π_p denote the halfplane the transition into which across the interval $[t_0, t_1]$ of the A -line increases the number of IPs. The following definition is important.

Definition 3.5. All planes in $\Pi_{pr}(x) := \Pi_p(x) \cap \Pi_*(x)$ are called proper, and all planes in $\Pi_{im}(x) := \Pi(x) \setminus (\Pi_{pr}(x) \cup \Omega_{crit}(x))$ are called improper.

4. ASSIGNING WEIGHTS

According to the general scheme of [Kat03], we have to assign weights to all IPs in $\Pi \cap C_{PI}(x)$ given any plane $\Pi \ni x$. In the first subsection we state a method of assigning weights to the IPs that ensures that the resulting reconstruction algorithm is exact and of the FBP type. The method of assigning weights depends not only on the IPs of Π with C inside $I_{PI}(x)$, but also on some extra IPs that are located outside $I_{PI}(x)$. This is the price we have to pay for considering a rather general class of curves. Additionally, the method depends on a function $q(\Pi)$. The way of choosing extra IPs is described in the second subsection, and the function $q(\Pi)$ is described in the third subsection.

Given a proper plane $\Pi \in \Pi_{pr}(x)$, $IP(x, \Pi)$ denotes a subset of $C \cap \Pi$ that includes all IPs located inside $I_{PI}(x)$ and, possibly, some extra IPs that are located outside of $I_{PI}(x)$. We assume that if $IP(x, \Pi)$ contains those extra IPs, then all of them are either below s_b or above s_t . Later we show that properties P1–P6 guarantee that the set $IP(x, \Pi)$ with the required properties can be chosen in an efficient manner.

4.1. General scheme. Consider a plane Π through x , which is transversal to $C_{PI}(x)$. The weights are assigned according to the diagrams below. We use the convention that ‘+’ represents weight 1, ‘-’ represents weight -1, and the i -th element in the diagram gives the weight of the i -th IP inside $I_{PI}(x)$.

W1. If Π is improper:

$$\boxed{+ \quad - \quad + \quad \dots \quad + \quad - \quad +}$$

W2. If Π is proper and there are no IPs in $IP(x, \Pi)$ outside $I_{PI}(x)$:

$$\boxed{\pm \quad + \quad - \quad \dots \quad - \quad + \quad \mp}$$

W3. If Π is proper and some IPs in $IP(x, \Pi)$ are greater than $s_t(x)$:

$$\boxed{+ \quad + \quad - \quad \dots \quad - \quad + \quad -}$$

W4. If Π is proper and some IPs in $IP(x, \Pi)$ are smaller than $s_b(x)$:

$$\boxed{- \quad + \quad - \quad \dots \quad - \quad + \quad +}$$

An addition to the rule W2, which explains how to assign ‘+’ and ‘-’ to the first and last IPs, is given below. Here we note only that whenever the first IP has weight ‘+’, then the last IP has weight ‘-’, and vice versa.

Fix some $s \in I_{PI}(x)$. As in [Kat03], define $\beta(s, x) = \frac{x-y(s)}{|x-y(s)|}$. Let θ be a polar angle in the plane $\beta^\perp(s, x)$. Then $\alpha(\theta) \in \beta^\perp(s, x)$ denotes the corresponding unit vector. The plane through x and $y(s)$ and perpendicular to $\alpha(\theta)$ is denoted by $\Pi(\theta)$. Without loss of generality we may assume that (a) $\Pi(\theta = 0)$ contains $L_{PI}(x)$, (b) sufficiently small $\theta < 0$ correspond to proper planes, and (c) sufficiently small $\theta > 0$ correspond to improper planes (see Fig. 8).

According to [Kat03], filtering planes are found by locating discontinuities of the function

$$(4.1) \quad \phi(s, x, \theta) = \text{sgn}(\alpha \cdot \tau(s))n(s, x, \alpha), \quad \alpha = \alpha(\theta) \in \beta^\perp(s, x),$$

where $n(s, x, \alpha)$ is the weight of $y(s)$ as an IP in $\Pi(\theta) \cap C_{PI}(x)$. The following result is the first step towards establishing the FBP structure of the resulting algorithm.

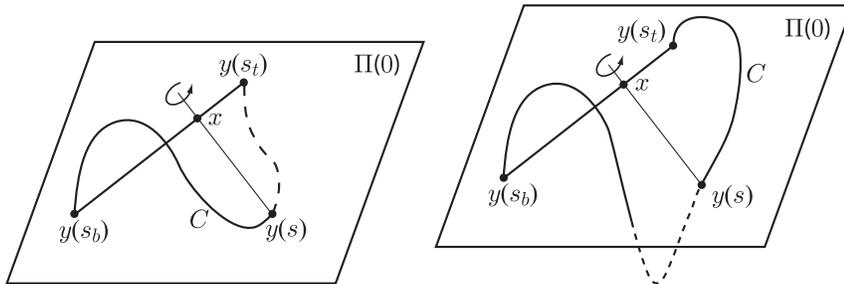


FIGURE 9. Case I (left panel) and case II (right panel) from the proof of proposition 4.1

Proposition 4.1. Fix any $x \in U$ and $s \in I_{PI}(x)$, $y(s) \neq y(s_b), y(s_t)$. Then $n(s, x, \alpha(\theta))$ is continuous across $\theta = 0$.

Proof. We write $\Pi(0^-)$ to denote (proper) planes $\Pi(\theta)$ for $\theta < 0$ sufficiently small, and $\Pi(0^+)$ to denote (improper) planes $\Pi(\theta)$ for $\theta > 0$ sufficiently small.

As it was mentioned in [Kat06] for the standard helix (see [Kat06], section 4), two cases are possible:

- I. Locally, the sections of $C_{PI}(x)$ attached to $y(s_b)$ and $y(s_t)$, respectively, are on different sides of $\Pi(0)$ (see Fig. 9, left panel).
- II. Locally, the sections of $C_{PI}(x)$ attached to $y(s_b)$ and $y(s_t)$, respectively, are on the same side of $\Pi(0)$ (see Fig. 9, right panel).

Case I corresponds to planes on the A -line that are located between t_0 and t_1 , whereas case II corresponds to all other planes in the intersection of the A -line with $\Pi_*(x)$ (see the discussion after proposition 3.4 and definition 3.5). In the first case as θ increases through 0, two IPs disappear at the endpoints of the PI line (see Figure 10). In the second case, when θ increases through 0, an IP disappears at one endpoint of $I_{PI}(x)$, and another IP appears at the other endpoint (see Fig. 12). As is known, only case I happens when the curve is a standard helix.

It now follows that we need to consider three cases:

- I. Both IPs disappear as θ increases through 0 (see Fig. 10). In this case s cannot be the first or the last IP of $\Pi(0^-)$. Since $\Pi(0^-)$ is a proper plane and $\Pi(0^+)$ is an improper plane, using rules W1 and W2-W4 gives that $n(s, x, \alpha(\theta))$ is continuous (see Fig. 11).

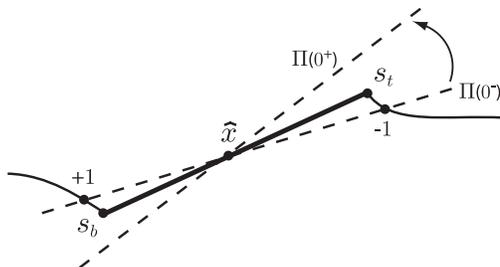


FIGURE 10. Projection onto the plane $(x - y(s))^\perp$. The case when two IPs disappear.

- II.1. The left-most IP disappears, while another IP emerges on the right (see Fig. 12, left panel). If s is neither the first nor the last IP of $\Pi(0^-)$, then, similarly to the preceding case, its weight does not change across $\theta = 0$ (cf. rules W1 and W3, W4 and Fig. 13).

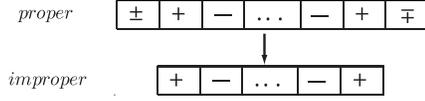


FIGURE 11. Change of weights when the first and last IPs disappear.

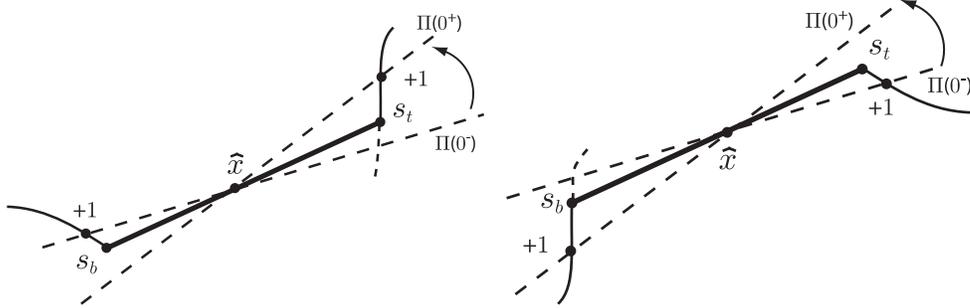


FIGURE 12. Projection onto the plane $(x - y(s))^\perp$. Left panel: an IP disappears on the left, right panel: an IP disappears on the right.

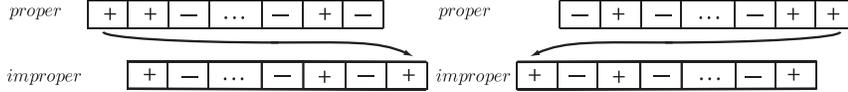


FIGURE 13. Change of weights when an IP disappears at one end and appears on the other end. Left panel - disappearance on the left, right panel - disappearance on the right.

Now consider the first and last IPs. Since the first IP disappears as θ increases, $y(s)$ can only be the last IP. Hence we need to show that the weight of the last IP does not change across $\theta = 0$, i.e. when $y(s)$ ceases to be the last IP. One sees that there are necessarily extra IPs of $\Pi(0^-)$ with C to the right of s_t (cf. Fig. 12, left panel). Also, due to the property that $IP(x, \Pi)$ contains IPs only on one side of $I_{PI}(x)$, there are no IPs to the left of s_b . According to rule W3, the weight of the last IP of $\Pi(0^-)$ with $C_{PI}(x)$ equals -1 . It now follows from Fig. 13, left panel, that the weight of the last IP is continuous as well.

II.2. The right-most IP disappears, while another IP emerges on the left (see Fig. 12, right panel). The proof is analogous to the previous case. See Fig. 13, right panel.

□

Remark 4.2. The preceding argument demonstrates the importance of the property that $IP(x, \Pi)$ contain IPs only on one side of $I_{PI}(x)$. If $\Pi(0^-)$ is such that an IP disappears at one end and appears on the other end across $\theta = 0$, then this property allows us to determine at which end the disappearance takes place. This, in turn, allows us to fix the weights so that $n(s, x, \alpha(\theta))$ is continuous across $\theta = 0$. If the extra IPs were allowed on both sides of $I_{PI}(x)$, there would be no easy way of finding out what is happening to IPs across $L_{PI}(x)$.

4.2. The set $IP(x, \Pi)$. In order to define $IP(x, \Pi)$, we need to use some results from section 3. We can use these results, because the PI lines introduced in section 2 satisfy the more general definition of PI lines given in section 3. We formally state this result as follows.

Proposition 4.3. *Let $H(s_b, s_t)$ be a chord of C , which is a PI line in the sense of definition 2.2. Then $H(s_b, s_t)$ is also a PI line in the sense of definition 3.1 for any $x \in H(s_b, s_t)$, where one can take $D = N_{max}(t)$ for any $t \in [q_{min}(s_t), s_b]$ or $D = N_{min}(t)$ for any $t \in [s_t, q_{max}(s_b)]$.*

Proof. Follows immediately from proposition 2.5. \square

We now introduce a convention for choosing normals to proper planes. Given $\Pi \in \Pi_{pr}(x)$, denote the IPs of Π with $C_{PI}(x)$ by $s_1 < s_2 < \dots < s_{2k+1}$. If C is non-smooth at one of the IPs, there can be infinitely many values of the parameter that correspond to the same IP. In the remainder of the paper for the most part we talk about points $y \in C$, so it is irrelevant which particular s is chosen to represent any $y \in C$ as long as $y(s) = y$. Whenever a statement does depend on which parameter value is chosen, it is understood that the statement holds for all equivalent values of the parameter. Recall that for any proper plane there are at least three IPs inside $I_{PI}(x)$. Choosing the normal as $\bar{N} := (y(s_3) - y(s_1)) \times (y(s_2) - y(s_1))$, we have

$$(4.2) \quad \bar{N} \cdot N_{max}(t) > 0 \text{ for any } t \in [q_{min}(s_{2k+1}), s_1], \quad \bar{N} \cdot N_{min}(t) < 0 \text{ for any } t \in [s_{2k+1}, q_{max}(s_1)].$$

This follows from (a) the convexity of the projections of $C(s, q_{max}(s))$ and $C(q_{min}(s), s)$ onto $N_{max}^\perp(s)$ and $N_{min}^\perp(s)$, respectively, and (b) the observation that by letting $y(s_1), y(s_2)$, and $y(s_3)$ approach each other while keeping $s_1 < s_2 < s_3$, we can get any osculating plane $\Pi_{osc}(q)$, $q \in [s_1, s_{2k+1}]$, by noting that in the process the dot product $\bar{N} \cdot N_{max}(t)$ is never zero and using P1. In what follows we always assume that normals \bar{N} to proper planes are chosen according to this convention.

For a generic $\Pi \in \Pi_{pr}(x)$ we denote the IPs of Π with the interior of $C(q_{min}(s_{2k+1}), s_b)$ by $l_1 < \dots < l_m$, and the IPs of Π with the interior of $C(s_t, q_{max}(s_1))$ – by $r_1 < \dots < r_n$. It is easy to see that this definition is unambiguous, i.e. it does not depend on the specific values of s_1 and s_{2k+1} (recall that we are using the non-standard parameterization). Indeed, pick any IP $r > s_t$. Suppose there are two equivalent values $s_1^{(1)}$ and $s_1^{(2)}$ such that $q_{max}(s_1^{(1)}) < r < q_{max}(s_1^{(2)})$. Using the monotonicity of q_{max} there is another equivalent value s_1^* such that $q_{max}(s_1^*) = r$, which contradicts (4.2). Using (4.2) again we immediately get

Proposition 4.4. *Pick $x \in U$ and a plane $\Pi \in \Pi_{pr}(x)$. Then $\bar{N} \cdot N_{max}(l_p) > 0$, $p = 1 \dots m$, if $m > 0$, and $\bar{N} \cdot N_{min}(r_p) < 0$, $p = 1 \dots n$, if $n > 0$.*

Consider $x \in U$ and $s \in I_{PI}(x)$, $y(s) \neq y(s_b), y(s_t)$. Suppose case II.1 happens (see the proof of proposition 4.1), i.e. an IP emerges across s_t and another IP disappears across s_b as one rotates the plane $\Pi(\theta)$ around $\beta(s, x)$. As we have already seen, $\Pi(0^-)$ has IPs with C above s_t . Moreover, there are necessarily IPs that are located *to the right of \hat{x}* on $DP_+(s)$. In the same way, if a new IP emerges on the left, then for small $\theta < 0$ there are necessarily IPs below s_b *to the left of \hat{x}* on $DP_-(s)$.

It was shown in section 3 that case II happens when the plane $\Pi(0)$, which corresponds to a point on the A -line in the plane $x_3 = 1$, is located either to the left of t_0 or to the right of t_1 . Next consider a proper plane Π close to $\Pi(0)$. Thus, on the plane $x_3 = 1$, Π corresponds to a point located close to the A -line. Suppose, for example, that the point is to the left of t_0 (see Fig. 14). In this case all B_s -curves through Π intersect the A -line to the left of t_0 except for the B_s -curve that corresponds to the last IP of Π with $C_{PI}(x)$. Indeed, in order to intersect the A -line to the right of t_0 , the B_s would need to have a slope arbitrarily close to 0, but only B_s with s close to s_b or s_t has this property. Hence, if Π is close to the A -line to the left of t_0 , then case II.2 necessarily

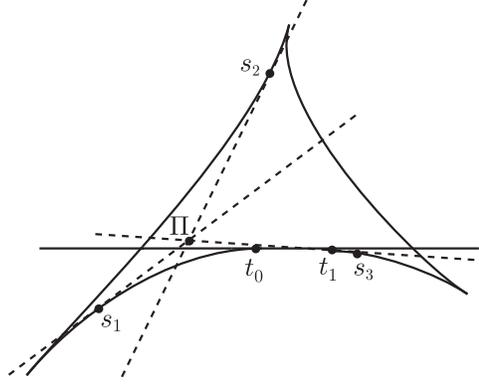


FIGURE 14. B_s -curves of a plane close to the A -line to the left of t_0

happens for all B_{s_j} -curves, $1 \leq j < 2k + 1$. In other words, all proper planes that are sufficiently close to the A -line to the left of t_0 have IPs below s_b that are located to the left of \hat{x} on $DP_-(s_j)$ for all $1 \leq j < 2k + 1$. Similarly, if a proper plane is close to the A -line to the right of t_1 , then case II.1 happens for all B_{s_j} -curves, $1 < j \leq 2k + 1$, i.e. there are IPs above s_t that are located to the right of \hat{x} on $DP_+(s_j)$ for all $1 < j \leq 2k + 1$.

Hence, to fix the weights of the IPs inside $I_{PI}(x)$ according to rules W3, W4, we do not have to pay attention to all the IPs $\Pi \cap C$ located close to $I_{PI}(x)$, but only to those IPs that are located on the appropriate side of Π relative to the line through x and $y(s_j)$ for some j , $1 < j < 2k + 1$. More precisely, if an IP is above s_t , it must be to the right of the line in order to be used in rule W3. Similarly, if an IP is below s_b , it must be to the left of the line in order to be used in rule W4. Choosing different j affects the algorithm. The simplest case occurs if we choose the middle IP inside $I_{PI}(x)$: $j = k + 1$. In particular, the set $IP(x, \Pi)$ can be defined in a simple way: $IP(x, \Pi)$ consists of the following IPs:

- (H1) All IPs l_p , $1 \leq p \leq m$;
- (H2) All IPs that belong to $I_{PI}(x)$;
- (H3) All IPs r_p , $1 \leq p \leq n$.

The following proposition says that if the set $IP(x, \Pi)$ is defined according to (H1)-(H3), then the extra IPs are indeed located in the appropriate halfplanes.

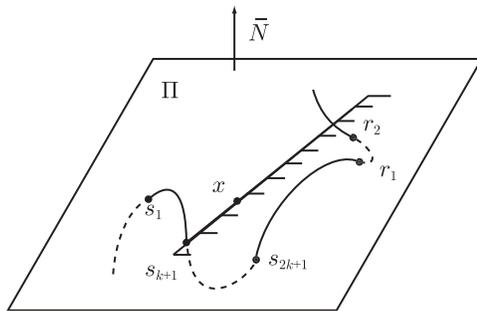
Proposition 4.5. *For all $x \in U$ and planes $\Pi \in \Pi_*(x)$ with the canonically chosen normal \bar{N} (i.e. with the normal chosen so that it satisfies (4.2)) the set $IP(x, \Pi)$ satisfies the following properties:*

- (I1) *For all p , $1 \leq p \leq m$ one has $[x - y(s_{k+1}), y(l_p) - y(s_{k+1}), \bar{N}] > 0$, i.e. all IPs $y(l_p)$ are located to the left of the line through x and $y(s_{k+1})$ on Π ;*
- (I2) *For all p , $1 \leq p \leq n$, one has $[x - y(s_{k+1}), y(r_p) - y(s_{k+1}), \bar{N}] < 0$, i.e. all IPs $y(r_p)$ are located to the right of the line through x and $y(s_{k+1})$ on Π .*

The proof of the proposition is given in appendix C. The statement of proposition 4.5 is illustrated in Fig. 15. It shows two IPs r_1, r_2 , which are located to the right of the line through x and $y(s_{k+1})$.

To summarize, the definition of $IP(x, \Pi)$ is very simple: we use those IPs outside $I_{PI}(x)$ that belong to $C(q_{min}(s_{2k+1}), q_{max}(s_1))$. Now we can formulate the main result of this section.

Proposition 4.6. *Fix any $x \in U$ and $\Pi \in \Pi_{pr}(x)$. Then $IP(x, \Pi)$ contains IPs either to the left or to the right of $I_{PI}(x)$.*


 FIGURE 15. Illustration of the properties of $IP(x, \Pi)$ (see proposition 4.5)

Proposition 4.6 says that the set $IP(x, \Pi)$ can indeed be used to assign weights according to rules W3, W4 in an unambiguous fashion. The proof is quite complicated and given in appendix C.

4.3. The function $q(\Pi)$. Now we discuss a way of choosing signs in the definition of weights of the first and last IPs in rule W2. Consider a function

$$(4.3) \quad q : \Pi_{pr}(x) \rightarrow \mathbb{R}.$$

Recall that $\Pi_{pr}(x)$ does not include planes through x that are tangent to $C_{PI}(x)$. Although $q(\Pi)$ is defined only for planes intersecting $C_{PI}(x)$ transversely, the function $q(\Pi)$ that we suggest below is actually rather well-behaved. In particular, for the most part $q(\Pi)$ is continuous across planes tangent to $C_{PI}(x)$. Define

$$(4.4) \quad q(\Pi) := \int_{s_1}^{s_{2k+1}} \bar{N} \cdot (y(t) - y(s_1)) |y'(t)| dt, \quad \Pi \in \Pi_{pr},$$

where \bar{N} is the normal to Π chosen according to property P1. The reason why $|y'(t)|$ is inserted in (4.4) is to make q independent of the parametrization. In particular, if $y(t) \equiv \text{const}$ over an interval, i.e. $|y'(t)| \equiv 0$, then the contribution of this interval into q equals zero. For the same reason, q does not depend on which particular values of the parameter are chosen for s_1 and s_{2k+1} . If C is the usual helix, then Π has only three IPs with $C_{PI}(x)$ and $q(\Pi)$ gives exactly the same filtering plane as the one determined by the function $\psi \equiv 1$ in [Kat04b]. Using (4.4), append rule W2 as follows:

W2'. If the plane Π is proper and $IP(x, \Pi) \subset I_{PI}(x)$, then the first (resp. last) IP gets weight 1 (resp. -1) if $q(\Pi) > 0$ and -1 (resp. 1) otherwise.

The function $q(\Pi)$ has some useful properties. From (4.4), $q(\Pi)$ is discontinuous only when Π touches $C_{PI}(x)$ below the first IP or above the last IP, because in this case s_1 or s_{2k+1} may jump. Additionally, $q(\Pi(\theta))$ is monotone on each interval where it is continuous.

Proposition 4.7. *Pick $x \in U$ and $s \in I_{PI}(x)$. Assume that either $s = s_1$ is the first IP or $s = s_{2k+1}$ is the last IP. Then when $\Pi(\theta) \in \Pi_{pr}(x)$ (i.e., $\theta < 0$), $q(\Pi(\theta))$ is continuous and monotonically decreasing as long as no new IPs appear below s_1 or above s_{2k+1} inside $I_{PI}(x)$.*

Proof. Similarly to section 3, project $C_{PI}(x)$ onto a plane parallel to $D = N_{max}(s_b)$ and $L_{PI}(x)$ and not containing x (see Fig. 16). We can still denote this plane $x_1 = 1$. Then the projected curve $\hat{C}_{PI}(x)$ approaches the asymptote L_b (resp., L_t) as $s \rightarrow s_b^+$ (resp., $s \rightarrow s_t^-$). Both L_b and L_t are parallel to $L_{PI}(x)$. Assume, for example, that L_b is above L_t . The other case is completely analogous. L_b cannot coincide with L_t , because property P3 implies $[\tau(s_b), \tau(s_t), y(s_t) - y(s_b)] \neq 0$.

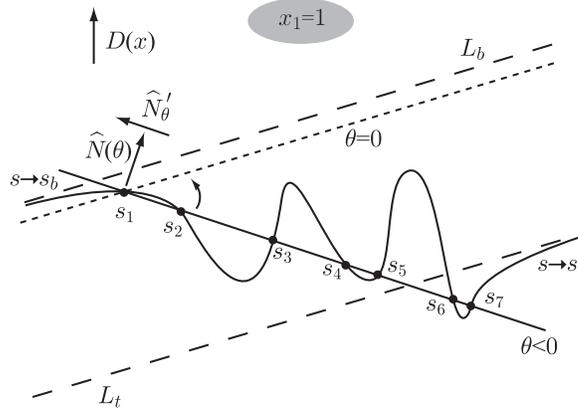


FIGURE 16. Monotonicity of $q(\Pi(\theta))$ when $\Pi(\theta)$ rotates around $\beta(s_1, x)$

Any plane through x and $y(s)$, $s \in (s_b, s_t)$, corresponds to a line through $\hat{y}(s)$ on the plane $x_1 = 1$. As is easily seen from Fig. 16, if $\hat{y}(s)$ is located between L_b and L_t , then the rotation of $\Pi(\theta)$ across $\theta = 0$ leads to a disappearance of two IPs. This shows that the direction of increasing θ corresponds to the counter-clockwise rotation in the plane.

Suppose we rotate $\Pi(\theta)$ around the first IP $y(s_1)$. The case of rotation about the last IP is analogous. It is clear that $s_1(\theta) = \text{const}$ and $s_{2k+1}(\theta)$ is a continuous function of θ as long as no new IPs appear below s_1 or above s_{2k+1} . By (4.4), $q(\Pi(\theta))$ is continuous under the same conditions.

Assuming that $\Pi = \Pi(\theta)$ in (4.4) and differentiating with respect to θ gives

$$(4.5) \quad q(\Pi(\theta))'_\theta = \int_{s_1}^{s_{2k+1}} \bar{N}'_\theta \cdot (y(t) - y(s_1)) |\dot{y}(t)| dt,$$

where we have used that $\bar{N} \cdot (y(s_{2k+1}) - y(s_1)) = 0$. Note that $\bar{N}'_\theta \cdot (y(t) - y(s_1)) \leq 0$, $s_1 \leq t \leq s_{2k+1}$. See Fig. 16, which shows the projections of $\bar{N}(\theta)$ and \bar{N}'_θ onto the plane $x_1 = 1$. Hence the right side of (4.5) is negative, and $q(\Pi(\theta))'_\theta < 0$. \square

Since the planes $\Pi \in \Pi_{pr}$ satisfying $q(\Pi) = 0$ form a boundary between the set of planes with different weights, we need to make sure that this boundary is a set of measure zero in S^2 . This question is addressed by the following proposition.

Proposition 4.8. *Denote the set of planes $\Pi \in \Pi_{pr}$ such that $q(\Pi(\theta)) = 0$ by Ω . Then Ω is a set of measure zero in S^2 .*

Proof. By Proposition 4.7, $q(\Pi)$ is smooth away from the T -curve and A -line and has nonzero gradient due to monotonicity. Hence, Ω has measure zero. \square

A major advantage of the function $q(\Pi)$ is a simple algorithm for finding the family of solutions to $q(\Pi) = 0$ that contain $y(s)$. Denote for simplicity $s = s_1$ and $s' = s_{2k+1}$. Introduce the vector

$$(4.6) \quad Y := \int_s^{s'} (y(t) - y(s)) |\dot{y}(t)| dt.$$

Then define $\bar{N} := Y \times (y(s') - y(s))$. Clearly, one has

$$\int_s^{s'} \bar{N} \cdot (y(t) - y(s)) |\dot{y}(t)| dt = \bar{N} \cdot Y = [Y, Y, y(s') - y(s)] = 0.$$

Fix $x \in U$ and $s \in I_{PI}(x)$. As follows from (4.1) and rule W2', a plane Π through x and $y(s)$ is a filtering plane if (a) $q(\Pi) = 0$, (b) s and s' are the first and last or the last and first IPs in $\Pi \cap C_{PI}(x)$, and (c) the set $IP(x, \Pi)$ contains no IPs outside $I_{PI}(x)$. Suppose we found $s' \in (s_{min}(s), s_{max}(s))$ such that the plane Π with normal \bar{N} constructed as above contains x . We call this Π a 'potential' filtering plane, because it does not necessarily satisfy requirements (b) and (c) stated above. For a fixed $s \in I_0$, all such potential filtering planes can be found by varying $s' \in (s_{min}(s), s_{max}(s))$ and using the discussion above.

5. THE ALGORITHM

In the first subsection we formulate the reconstruction algorithm that results from the weight function described in section 4. Then we prove the main result of the paper, which says that the algorithm is of the FBP type.

5.1. Inversion formula. Let $D_f(s, \Theta) = \int_0^\infty f(y(s) + t\Theta)dt$, $|\Theta| = 1$, denote the cone beam transform of f .

Theorem 5.1. *Let C be a curve (2.1), which satisfies conditions P1–P6. Let $I_0 \subset I$ be an interval, such that the set U defined by (2.11) is non-empty. For any $f \in C_0^\infty(U)$ and $x \in U$ which admits a PI line with $I_{PI}(x) \subset I_0$ one has*

$$(5.1) \quad f(x) = -\frac{1}{4\pi^2} \int_{I_{PI}(x)} \frac{1}{|x - y(s)|} \sum_{k=1}^{N(s,x)} c_k \int_0^{2\pi} \frac{\partial}{\partial q} D_f(q, \Theta_k(s, x, \gamma)) \Big|_{q=s} \frac{d\gamma}{\sin \gamma} ds.$$

Here

$$\begin{aligned} \beta(s, x) &:= (x - y(s))/|x - y(s)|, \quad e_k(s, x) := \beta(s, x) \times u_k(s, x), \\ \Theta_k(s, x, \gamma) &:= \cos \gamma \beta(s, x) + \sin \gamma e_k(s, x), \end{aligned}$$

$\theta_1, \dots, \theta_{N(s,x)}$ are the points where $\phi(s, x, \theta)$ is discontinuous, $c_k := \phi(s, x, \theta_k^+) - \phi(s, x, \theta_k^-)$ are the values of the jumps, and $u_k(s, x)$ are the normal vectors to the planes $\Pi(\theta_k)$ chosen according to our convention (cf. Property P1 and (4.2)).

Proof. At the end of section 5.2 we show that for all $x \in U$ and $s \in I_{PI}(x)$ the number of discontinuities of $\phi(s, x, \theta)$ is finite. The rest follows from inversion formula (2.28) in [Kat03]. \square

Note that the nonstandard parametrization (2.1), (A1)–(A3) of C does not affect the validity of (5.1). Indeed, if $y(s) \equiv \text{const}$ over an interval, then $\frac{\partial}{\partial q} D_f(q, \cdot) \Big|_{q=s} \equiv 0$ for all s in the interval.

5.2. FBP structure of the algorithm. We now show that the resulting algorithm is of the FBP type and uses three one-dimensional families of filtering lines. To do that, we need to consider the discontinuities of

$$(5.2) \quad \phi(s, x, \theta) = \text{sgn}(\alpha \cdot \tau(s)) n(s, x, \alpha), \quad \alpha = \alpha(\theta) \in \beta^\perp(s, x),$$

where $n(s, x, \alpha)$ is the weight of $y(s)$ as an IP in $\Pi(\theta) \cap C_{PI}(x)$. The function $\phi(s, x, \theta)$ can be discontinuous when either $\text{sgn}(\alpha \cdot \tau(s))$ or $n(s, x, \alpha)$ is discontinuous.

Discontinuities caused by the first factor in (5.2) may cause planes tangent to C at $y(s)$ to be filtering planes. These planes pass through $y(s)$ and are parallel to $\dot{y}(s)$, so the family, which we denote \mathcal{L}_1 , is one-parametric.

By proposition 4.1, $n(s, x, \alpha(\theta))$ is continuous across the PI line of x . Assuming $\Pi(\theta)$ does not intersect $L_{PI}(x)$, $n(s, x, \alpha(\theta))$ can be discontinuous at some θ_0 only for the following reasons:

- (J1) $y(s)$ ceases to be the first/last IP with $C_{PI}(x)$; or
- (J2) $q(\Pi(\theta_0^-)) \leq 0 \leq q(\Pi(\theta_0^+))$; or

(J3) an IP outside of $I_{PI}(x)$ enters/leaves the set $IP(x, \Pi)$.

Case (J1) corresponds to a filtering plane that passes through $y(s)$ and touches $C_{PI}(x)$, i.e. (J1) leads to a one-dimensional family of filtering planes, denoted \mathcal{L}_2 . (J2) actually describes two cases. If $q(\Pi(\theta))$ is discontinuous at θ_0 , then $\Pi(\theta_0)$ is tangent to $C_{PI}(x)$, which gives the same family as in (J1). If $q(\Pi(\theta))$ is continuous at θ_0 and $q(\Pi(\theta_0)) = 0$, then $\Pi(\theta_0)$ belongs to the family of potential filtering planes through $y(s)$, which is one-dimensional (cf. section 4.3). Denote it \mathcal{L}_3 . Finally, we need to show that (J3) leads to a one-dimensional family of filtering planes. This statement is contained in the following proposition.

Proposition 5.2. *If the number of IPs in $IP(x, \Pi(\theta))$ on either side of $I_{PI}(x)$ changes across $\theta = \theta_0$, then $\Pi(\theta_0)$ is tangent to $C(q_{min}(s), q_{max}(s))$.*

Proof. Denote the IP that is leaving $IP(x, \Pi(\theta))$ by $r(\theta)$. Suppose $r(\theta)$ is above $I_{PI}(x)$ (the other case can be handled analogously). Denote the IPs of $\Pi(\theta)$ with $C_{PI}(x)$ for θ sufficiently close to θ_0 by $s_1(\theta) < \dots < s_{2k+1}(\theta)$. By definition, a point can leave $IP(x, \Pi(\theta))$ either when r disappears through tangency or when $r = q_{max}(s_1)$ when $\theta = \theta_0$. The latter possibility contradicts (4.2) with $t = s_1(\theta_0)$. \square

Hence, IPs outside $I_{PI}(x)$ can leave $IP(x, \Pi)$ only through tangency, so (J3) also leads to the family \mathcal{L}_2 .

Fix any $y(s) \in C$ and $x \in U$ such that $y(s) \in I_{PI}(x)$. Even though there are only three families of filtering planes, there can be more than one plane from the same family passing through x . We show that this number is finite. It is clear that there can be at most one plane through x from \mathcal{L}_1 . Condition P6 gives that the number of planes through x from \mathcal{L}_2 is uniformly bounded. From Proposition 4.7, $q(\Pi(\theta))$ can be discontinuous only when $\Pi(\theta)$ is tangent to $C(q_{min}(s), q_{max}(s))$. Because of the monotonicity of $q(\Pi(\theta))$, there can be at most one root to $q(\Pi(\theta)) = 0$ inside each interval where q is continuous. Hence the number of filtering planes from \mathcal{L}_3 is uniformly bounded as well.

6. NUMERICAL EXPERIMENTS

Two numerical experiments were conducted to validate our results. In both of them the detector was flat.

In the first experiment we used the standard clock phantom, and the source trajectory was the variable radius/variable pitch helix given by

$$y(s) = \left(R(s) \cos(s), R(s) \sin(s), \frac{h_0}{2\pi} (s + 1.95 \sin(1.2s)) \right),$$

where $R(s) = R_0(1 + 0.15 \sin(s/3))$, $R_0 = 600$ mm and $h_0 = 35$ mm. The horizontal slice $x_3 = 0$ was reconstructed inside the disk

$$D = \{(x_1, x_2, 0) | x_1^2 + x_2^2 \leq (0.4R_0)^2\}.$$

It was numerically verified that the trajectory satisfies properties P1–P6. In particular, in order to show that reconstruction is possible inside D , we verified numerically that

$$D \subset \bigcap_{s \in [q_{min}(0), q_{max}(0)]} Cy'_{min}(s) \cap Cy'_{max}(s).$$

Since $y(s)$ is below the plane $x_3 = 0$ for all $s < 0$ and above the plane $x_3 = 0$ for all $s > 0$, this ensures that each $x \in D$ has a unique PI line with both endpoints inside $[q_{min}(0), q_{max}(0)]$. Hence, $x \in Cy'_{min}(s) \cap Cy'_{max}(s)$ for all $s \in I_{PI}(x)$, and condition P5 is satisfied (where U is a sufficiently small neighborhood of D).

Parameter	Value	Units
Number of detector columns	1351	
Detector pixel size at isocenter	0.5×0.5	mm^2
Number of detector rows	151	
Views per rotation	1000	

TABLE 1. Simulation parameters for the first experiment

Parameters of the reconstruction are summarized in Table 1. The results of the reconstruction are shown in Figure 17.

In the second experiment we used a variable-pitch helix, which violates the convexity condition. The goal was to demonstrate that artifacts do appear if the algorithm of [KBH04] is used for reconstruction, and that the proposed algorithm reconstructs exactly in this case. Since the artifacts are expected to appear close to sharp x_3 -transitions in the phantom, we flattened the clock phantom by a factor of three in the x_3 -direction and decreased the vertical displacement between adjacent ellipsoids (unfortunately, this also led to more pronounced discretization artifacts). Since the convexity condition is violated, the algorithm of [KBH04] is not applicable for exact reconstruction. In order to apply it inexactly, we made the decision to use the filtering line of the smallest slope when more than one filtering line exists for a given pixel on the detector.

Parameters of our reconstructions are summarized in Table 2. The curve used is a variable-pitch helix that violates the convexity condition in the same way as in [KK06], although the perturbation is considerably more severe in this case. We have

$$y(s) = (R_0 \cos(s), R_0 \sin(s), \psi(s)),$$

where the x_3 -displacement $\psi(s)$ satisfies $\dot{\psi}(s) + \ddot{\psi}(s) = \mu(s)$. The following sample $\mu(s)$ was chosen for the simulations:

$$(6.1) \quad \mu(s) = \begin{cases} 8.333, & s < -0.7 \\ 37.5, & -0.7 < s \leq -0.4 \\ -10.750, & -0.4 \leq s \leq 0.4 \\ 37.5, & 0.4 \leq s < 0.7 \\ 8.333, & s \geq 0.7 \end{cases}$$

The numbers in the definition of μ are not round, because the perturbation of a constant speed helix had to be compactly supported. Since some of the conditions P1–P6 can be easily checked analytically for this trajectory, we discuss them separately. First, we have $N_{max}(s) = (0, 0, l_1(s)), l_1(s) > 0$ and $N_{min}(s) = (0, 0, l_2(s)), l_2(s) < 0$ for all s . It can be easily verified that the osculating plane never becomes vertical for the constant radius variable pitch helix, so P1 holds. Since the projection of $C(s, q_{max}(s))$ along $N_{max}(s)$ is always a circle, we have that the tangent line to C at $y(s)$ never intersects $\hat{C}(q_{min}(s), q_{max}(s))$, so P2 follows. It was checked numerically that C has no self-intersections, which verifies P3. Since the planes $\Pi_{min}(s)$ and $\Pi_{max}(s)$ are identical for all s , P4 is trivial. In order to verify P5, we verified numerically that there are no critical chords greater in angular length than $2 \arccos 0.4 \approx 2.31856$. Hence, reconstruction is possible inside

$$(6.2) \quad U := \{(x_1, x_2, x_3) | x_1^2 + x_2^2 < (0.4R_0)^2\}.$$

This also implies that PI lines exist for all $x \in U$ and are unique. Finally, since the torsion becomes negative over a single interval $[-0.4, 0.4]$ (see the definition (6.1)), we conclude that planes through $x \in U$ can have at most five IPs with $C_{PI}(x)$, so P6 also holds.

We need to note here that the violation of the convexity condition is such that for some s the projection of C onto the detector plane $DP(s)$ is not convex inside the field of view. The

Parameter	Value	Units
Source to rotation axis distance	600	mm
Number of detector columns	1351	
Detector pixel size at isocenter	0.5×0.5	mm^2
Number of detector rows	271	
Views per rotation	1000	

TABLE 2. Simulation parameters for the second experiment

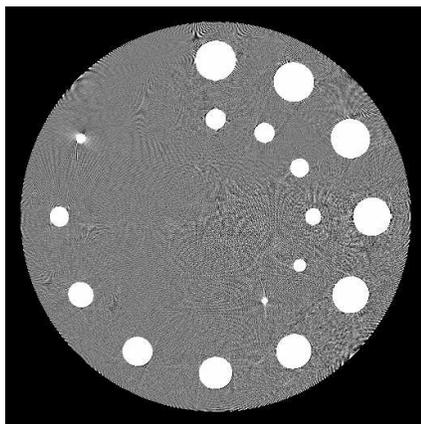


FIGURE 17. Reconstruction results for the first experiment

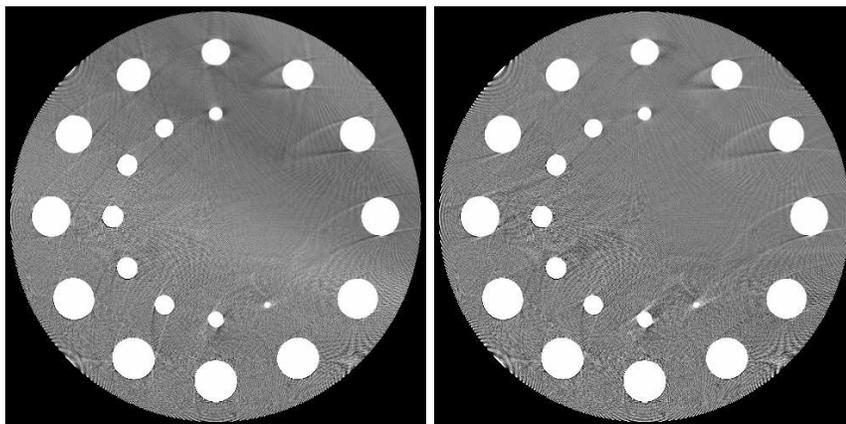


FIGURE 18. Reconstruction results for the second experiment: standard algorithm (left panel) and exact algorithm (right panel)

reconstruction results are shown in Figure 18. One sees that low-frequency artifacts appear in the reconstructed image when the original algorithm is used. No such artifacts appear in the image obtained with the proposed algorithm.

7. A SHORT SUMMARY

Here we summarize the main ideas and results of the paper.

- (K1) Introduced is a class of curves, which is an extension of the class of smooth curves with positive torsion proposed in [KK07]. Since torsion depends on the third derivative of the curve, violations of the positive torsion condition are likely to occur in practice. Points of non-smoothness are allowed on C , and torsion may become negative over some local intervals;
- (K2) A geometric condition is found, which controls how strongly the positivity of torsion is violated. The strength of a violation is measured by lengths of local critical chords and how close they come to the domain where reconstruction is performed;
- (K3) A geometric condition is found, which controls how the curve bends between neighboring turns. Loosely speaking, we assume that given any point $y(s) \in C$, all osculating planes $\Pi_{osc}(t)$ for t sufficiently far from s do not contain $y(s)$;
- (K4) Developed is a geometric construction that captures the entire dynamics of the Crofton symbol on the unit sphere. Originally this construction was proposed in [Kat06], but here it is developed to a greater extent. In the case when a curve segment admits a PI line, obtained are some geometric restrictions on the behavior of the boundaries across which the Crofton symbol is discontinuous. This leads to explicit necessary and sufficient conditions that describe those boundaries;
- (K5) The geometric construction mentioned above is used to develop a novel scheme for assigning weights to IPs between the Radon planes through x and $C_{PI}(x)$ (see [Kat06] for the first application of the construction to finding a weight function). The scheme uses IPs not only inside $I_{PI}(x)$, but also outside $I_{PI}(x)$. This appears to be a necessary condition if one wants to develop an efficient FBP inversion formula that works for sufficiently complicated curves C . Also a useful classification of the Radon planes into proper and improper is made;
- (K6) A deeper insight into the nature and role of the ψ function of [Kat04b, Kat02] is obtained, and its efficient generalization to other curves is proposed;
- (K7) Finally, the above items are combined to create an efficient cone beam transform inversion formula of the FBP type that applies to a general class of curves.

Let us elaborate on item (K6). In order to reduce detector requirements and make filtering lines as close to horizontal as possible, the weights of the first and last IPs in $\Pi \cap C_{PI}(x)$ must change across some curve γ on the unit sphere. More precisely, one is interested only in the section of γ located inside the domain of proper planes. The change, i.e. a discontinuity of the weight, results in a filtering plane. Since points on the unit sphere are identified with planes through x , the curve γ actually describes a family of filtering planes through x . Hence γ should be independent of x (in order to maintain the FBP structure), and depend only on the Radon plane. The equation of γ is therefore of the type $q(\Pi) = 0$, where q is a function defined on the set of proper planes. Given a reconstruction point x and a source position $y(s) \in C_{PI}(x)$, finding a filtering plane is equivalent to solving the equation

$$(7.1) \quad q(\Pi) = 0, \quad x, y(s) \in \Pi.$$

The function q of [Kat04b, Kat02] (where it is expressed in terms of the ψ function, cf. (5) in [Kat02] and (2.6) in [Kat04b]) explicitly uses all IPs in $\Pi \cap C_{PI}(x)$. For simple curves (e.g. helices), there can be at most three IPs, so solving (7.1) is fairly easy. When C is complicated and there can be more than three IPs, using all of them to construct q leads to equations (7.1) that are impractical to solve. The function q proposed here uses a different principle and solutions to (7.1) can be found very efficiently even when C is complicated.

The summary presented above shows that the results obtained in this paper contribute to a better understanding of the principles on which efficient cone beam inversion formulas are based. They can be used for the development of new inversion formulas for other classes of curves and also for the theoretical analysis of other aspects of the cone beam transform inversion.

APPENDIX A. PROOF OF PROPOSITION 2.5

Project $C(s, q_{max}(s))$ along $N_{max}(s)$ onto the plane $N_{max}^\perp(s)$. Denote the projected curve by $\hat{C}(s, q_{max}(s))$. To show that $\hat{C}(s, q_{max}(s))$ is convex, it suffices to notice that up to a positive factor the curvature of $\hat{C}(s, q_{max}(s))$ at $\hat{y}(t)$ is given by $[N_{max}(s), \tau(t), \dot{\tau}(t)]$, which is positive by property P1 for all $t \in (s, q_{max}(s))$ where C is smooth. When C is not smooth, the convexity is guaranteed by the extended definition of the osculating plane (see the paragraph following (A3)).

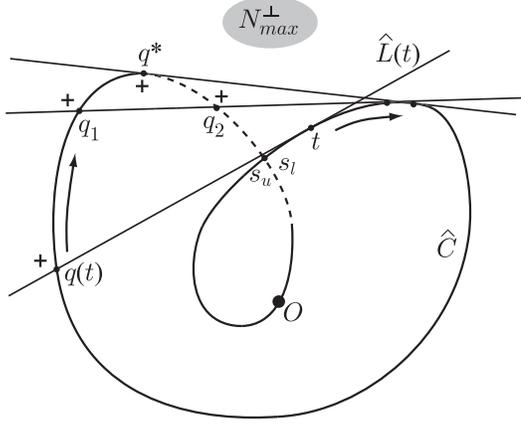
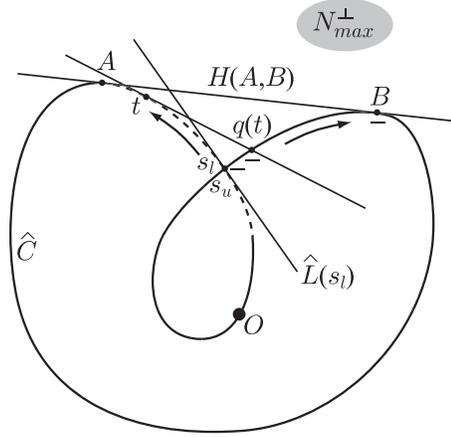
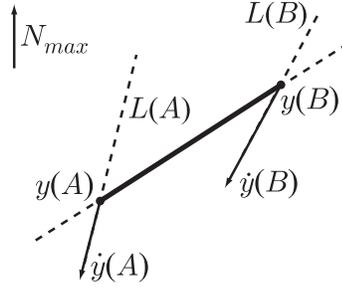


FIGURE 19. Projection onto the plane $N_{max}(s)^\perp$ in the case $q_{max}(s) < b$

The key step is to show that $\hat{C}(s, q_{max}(s))$ has no self-intersections. As in [KK07], we denote the projections of $y(s)$ and $y(q_{max}(s))$ by O . Even though $C(s, q_{max}(s))$ does not have self-intersections by property P3, $\hat{C}(s, q_{max}(s))$ may still self-intersect. For each such self-intersection we can mark which section of the curve lies above and which one lies below. This is illustrated by Fig. 19, where the lower section is denoted by a dashed curve. For a given $t \in [s, q_{max}(s)]$, let $L(t)$ denote the line tangent to C at $y(t)$. The projection of $L(t)$ onto $N_{max}^\perp(s)$ is denoted $\hat{L}(t)$. Thus $\hat{L}(t)$ is tangent to $\hat{C}(s, q_{max}(s))$ at $\hat{y}(t)$. If $\hat{L}(t)$ intersects \hat{C} at a point $\hat{y}(q)$, $q \neq t$, we can mark whether $L(t)$ is above (+) or below (-) $y(q)$, because $L(t)$ does not intersect $C(s, q_{max}(s))$ by condition P2. By the same condition, $L(t)$ does not intersect the maximal chord $H(s, q_{max}(s))$. This ensures that the ‘ \pm ’ signs do not change when the IP of \hat{L} and \hat{C} passes through O . In this derivation, as usual, we follow the convention that at the points where C is not smooth the tangent line rotates according to (2.2) sweeping an angle less than π by P1 and without intersecting $C(s, q_{max}(s))$ and $H(s, q_{max}(s))$ by P2. By condition P2, we can move the point of tangency $y(t)$ along C and the ‘ \pm ’ signs at the intersection points $\hat{y}(q)$ (if there are several of them) will not change (see Fig. 19). Hence, we can move the point of tangency along \hat{C} as long as $y(t)$ does not pass through O . Clearly, if two IPs $\hat{y}(q_k)$ and $\hat{y}(q_{k+1})$ approach each other at a point of tangency q^* , $q^* \neq t$, they should have the same signs. Hence, ‘ \pm ’ signs at the IPs of $\hat{L}(t)$ with $\hat{C}(s, q_{max}(s))$ propagate from one side of a tangency to the other (see Fig. 19, where the ‘+’ sign at q_1 can be propagated to q_2 via the point of tangency q^*).


 FIGURE 20. Projection onto the plane $N_{max}(s)^\perp$ in the case $q_{max}(s) < b$

 FIGURE 21. The critical chord $H(A, B)$ in the plane containing N_{max}

At first we assume that all self-intersections of \hat{C} are transversal. Consider the basic case that will be used throughout the rest of the proof. Pick a self-intersection of \hat{C} . Suppose that the lower point is $y(s_l)$, and the upper point is $y(s_u)$, i.e. $\hat{y}(s_l) = \hat{y}(s_u)$ and $(y(s_u) - y(s_l)) \cdot N_{max}(s) > 0$ (see Fig. 20). The idea is to associate a critical chord $H(A, B)$ with each self-intersection. By construction, $\hat{L}(s_l)$ intersects \hat{C} at s_u . As $\hat{L}(t)$ moves away from the self-intersection, the corresponding IP $\hat{y}(q), q = q(t)$, moves away too because of the convexity of \hat{C} . We move t in the direction, which guarantees that the angle between $\hat{\tau}(t)$ and $\hat{\tau}(q(t))$ decreases (see Fig. 20). Because of the convexity there exists a point $t = A$, which is the closest to s_l , such that $\hat{\tau}(A)$ and $\hat{\tau}(B)$, $B = q(A)$, have the same direction. Then the chord $H(A, B)$ is tangent to \hat{C} at two points. This gives $Q(A, B) = 0$, i.e. the chord is critical. In a similar fashion, moving $L(t)$ from s_u towards B we recover the same critical chord $H(A, B)$. By construction, $L(s_l)$ is below s_u , so we put a '-' next to s_u and keep track of the '-' sign as t moves from s_l to A and $q(t)$ moves from s_u towards B (see Fig. 20). Assuming that $y(t)$ does not pass through O , the '-' sign arrives at B when $t = A$, i.e. $L(A)$ passes below $y(B)$. In a similar fashion, $L(s_u)$ is above s_l , so we put a '+' next to s_l and keep track of the '+' sign as it moves towards A . Thus $L(B)$ passes above $y(A)$, assuming $y(t)$ does not pass through O on the way from s_u towards B . Our argument implies that the tangent vectors at A and B both point into the same half-plane relative to $H(A, B)$ in the plane containing $N_{max}(s)$ and passing through $H(A, B)$ (see Fig. 21). Hence $H(A, B)$ is a global critical chord, which contradicts proposition 2.4, because $[A, B]$ is inside (s, q_{max}) .

In the preceding argument it was essential that the point of tangency $y(t)$ does not pass through O as it moves from $y(s_l)$ towards $y(A)$ and from $y(s_u)$ towards $y(B)$. If this is not the case, we still get the same critical chord $H(A, B)$, but there is no guarantee that it is global. To describe this situation we say informally that O is between A and B . Hence, if we can find a self-intersection of \hat{C} such that O is not between A and B , then we obtain a global critical chord and, therefore, a contradiction. We now start with a self-intersection of \hat{C} such that O is between A and B , and consider possible behaviors of \hat{C} beyond B . More precisely, we consider possible locations of subsequent self-intersections.

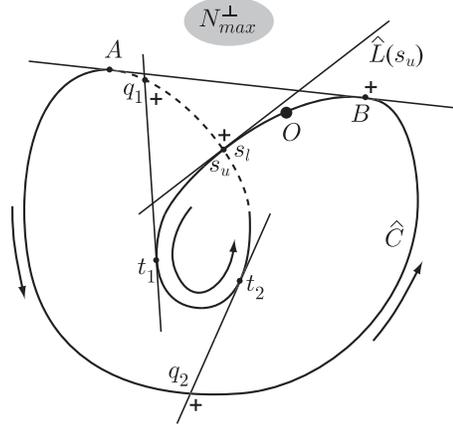


FIGURE 22. Projection onto the plane $N_{max}(s)^\perp$ in the case $q_{max}(s) < b$

Case 1. Suppose there are no more self-intersections, i.e. the curve closes at A after the only self-intersection (see Fig. 22). Suppose that O is located on the segment $C(s_u, B)$ (the other case can be handled similarly). We can no longer apply the same logic as before, since when $L(t)$ moves past O , some ‘ \pm ’ signs at the IPs can change. However, the following construction shows that this situation is impossible under our assumptions. Indeed, consider $L(s_u)$. Recall that there is a ‘+’ sign next to $y(s_l)$. Now we move the point of tangency from $t = s_u$ towards $t = A$ via the inner loop. Since $\hat{y}(t)$ does not pass O , the ‘+’ sign does not change (see Fig. 22). When the tangent is at A , the ‘+’ sign arrives at B , i.e. $L(A)$ is above $y(B)$. Starting now with $L(s_l)$ and moving the point of tangency towards A as in the basic case, we get that $L(A)$ is below $y(B)$. Thus $L(A)$ intersects $C(s, q_{max}(s))$, which contradicts P2.

Case 2. Suppose there is another self-intersection. Let $H(A', B')$ be the critical chord associated with it. Obviously, O must be between A' and B' . Otherwise we get a contradiction using the basic case. It is easy to see that the only way to avoid the basic case is to have $m \geq 1$ loops that wrap around the innermost loop to the right of O and $n \geq 1$ loops that wrap around the innermost loop to the left of O . See Fig. 23, which illustrates the case of $m = 3$ loops to the right of O and $n = 2$ loops to the left of O . Moreover, the line tangent to C at O must intersect all the loops (except the two innermost ones) in order to avoid the basic case. To establish a contradiction, we start with the innermost self-intersection to the right of O . Denote the lower point by q_l^1 and the upper point by q_u^1 . Suppose, for example, that q_l^1 is located on the segment of C that connects directly to O . The other case can be considered analogously. First, consider $L(t), t = q_u^1$. Since $y(q_u^1)$ is above $y(q_l^1)$, we put a ‘+’ sign near $\hat{y}(q_l^1)$. Next, move the point of tangency t in the direction away from the innermost loop. The ‘+’ sign moves away from the first self-intersection toward the second one. Let t^* be the point such that $\hat{L}(t^*)$ passes through the

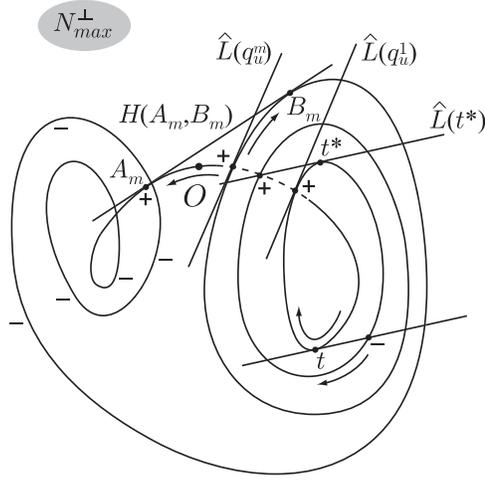


FIGURE 23. Multiple loops

second self-intersection (see Fig. 23). Because of the ‘+’ sign we get that the section of C located between O and q_l^1 in a neighborhood of the second self-intersection is located below $\hat{L}(t^*)$. Now consider $L(t), t = q_l^1$. Since $y(q_l^1)$ is below $y(q_u^1)$, we put a ‘-’ sign there. This time we move the point of tangency t in the direction away from O through the innermost loop. When t reaches t^* , the ‘-’ sign reaches the second self-intersection. This way we get that the other section of C in a neighborhood of the second self-intersection is located above $\hat{L}(t^*)$. Hence the section of C directly connecting O to q_l^1 is located below the other section of C in a neighborhood of the second self-intersection. By continuing in the same fashion we get that at all the self-intersections to the right of O , the section of C connecting O to q_l^1 is located below the other sections of C .

Pick now the outermost self-intersection, which is to the right of O . Let $H(A_m, B_m)$ be the associated critical chord. Consider $L(t), t = q_u^m$. Since $y(q_u^m)$ is above $y(q_l^m)$, we put a ‘+’ sign near $\hat{y}(q_l^m)$. Move the point of tangency t in the direction towards B_m (see Fig. 23). This ensures that the ‘+’ sign moves towards the left set of loops. When t reaches B_m , we get that $L(B_m)$ is located above $y(A_m)$. Next, consider the innermost self-intersection and $L(t), t = q_l^1$. Since $y(q_l^1)$ is below $y(q_u^1)$, we put a ‘-’ sign near $\hat{y}(q_u^1)$. By moving t towards B_m and going through all the loops to the right of O , we move the ‘-’ sign to the left loops. As was mentioned above, the \pm signs propagate through the points of tangency. So, if we move t back and forth over the outermost loop, we can obtain the ‘-’ sign at the point A_m when $t = B_m$. Since we cannot have the ‘+’ and ‘-’ signs at the same point, the proof is finished.

Suppose now that there can be tangential self-intersections. They can be of two types. If $\hat{y}(s_l) = \hat{y}(s_u)$ and $\hat{\tau}(s_l) \cdot \hat{\tau}(s_u) < 0$, then we deal with this case in the same way as with transversal self-intersections. If $\hat{y}(s_l) = \hat{y}(s_u)$ and $\hat{\tau}(s_l) \cdot \hat{\tau}(s_u) > 0$, then $H(s_l, s_u)$ is a global critical chord inside $(s, q_{max}(s))$.

We have shown that $\hat{C}(s, q_{max}(s))$ has no self-intersections. Thus the only remaining way for $\hat{C}(s, q_{max}(s))$ to be non-convex is to have an indentation with vertex at O . The shape of \hat{C} would be analogous to the one shown in Figure 20 with the point O located at the self-intersection and the inner loop removed. Then the basic case immediately gives a contradiction.

APPENDIX B. PROOFS OF PROPOSITIONS 3.3 AND 3.4

Proof of Proposition 3.3. We first prove the necessity. Suppose $C(s_b, s_t)$ is the PI segment of x , which is at the origin. We need to verify properties (F1)–(F3) for $u(s) := y(s) \times \tau(s) / |y(s) \times \tau(s)|$. Clearly, $u(s)$ is a continuous piecewise C^∞ curve. Property (F1) follows immediately from (D1). To verify (F2), note that the great circle C_A is given by $y(s_b)^\perp = y(s_t)^\perp$. Indeed, $y(s_b) \cdot u(s_b) = [y(s_b), y(s_b), \tau(s_b)] = y(s_t) \cdot u(s_t) = [y(s_t), y(s_t), \tau(s_t)] = 0$, so $u(s)$ starts and ends on C_A . Moreover,

$$y(s_b) \cdot \dot{u}(s_b^+) = [y(s_b), y(s_b), \ddot{y}(s_b^+)] = 0, \quad y(s_t) \cdot \dot{u}(s_t^-) = [y(s_t), y(s_t), \ddot{y}(s_t^-)] = 0,$$

which implies that $u(s)$ is tangent to C_A at $s = s_b$ and $s = s_t$. As was discussed prior to proposition 3.3, intersection of the plane $x_3 = 1$ with the plane through x and $\hat{u}(s)$ and perpendicular to $y(s)$ gives the tangent line to $\hat{u}(s)$ (even at points of nonsmoothness). By assumption, $y(s)$ is never parallel to D , and the projection of $y(s)$ onto $x_3 = 1$ given by $y(s) - D(y(s) \cdot D)$ rotates in the counterclockwise direction and sweeps π radians from $s = s_b$ to $s = s_t$. As is easily checked, the tangent line at $\hat{u}(s)$ and the projection of $y(s)$ are perpendicular, so the tangent line also rotates in the counterclockwise direction and sweeps π radians. This completes the proof of the necessity.

It is left to show the sufficiency. Suppose $u(s)$, $s \in [a, b]$, is some parametrization of the curve T . Using steps (E1)–(E3), find some curve $\tilde{y}(s) \in S^2$, $\tilde{y}_1(s) \geq 0$. From the involutivity of the Legendre transform, $u(s)$ is the T -curve of $\tilde{y}(s)$ at every point where both are smooth. Augment the set of tangent lines to $\tilde{y}(s)$ at points of nonsmoothness following (A3) and (2.2). Recall that the set of tangent lines to $\hat{u}(s)$ at points of nonsmoothness is augmented using the convention stated after (E3). Then $u(s)$ is the T -curve of $\tilde{y}(s)$. Since (a) $u(s)$ is tangent to C_A at $s = a$ and $s = b$, and (b) the tangent lines to $\hat{u}(s)$ rotates monotonically and sweeps π radians, we get that $y(a) = -y(b)$ and the projection of $y(s)$ onto the plane $x_3 = 1$ rotates monotonically and sweeps π radians. \square

Proof of Proposition 3.4. From the proof of proposition 3.3 it follows that any line tangent to the T -curve of x corresponds to some $y(s) \in C(s_b, s_t)$. Similarly, any $y(s) \in C(s_b, s_t)$ gives a line tangent to the T -curve at s . Indeed, if u smooth at s , this follows directly from (3.2). At a point where u is not smooth, this follows from our extended definition of a tangent line. This proves (G1).

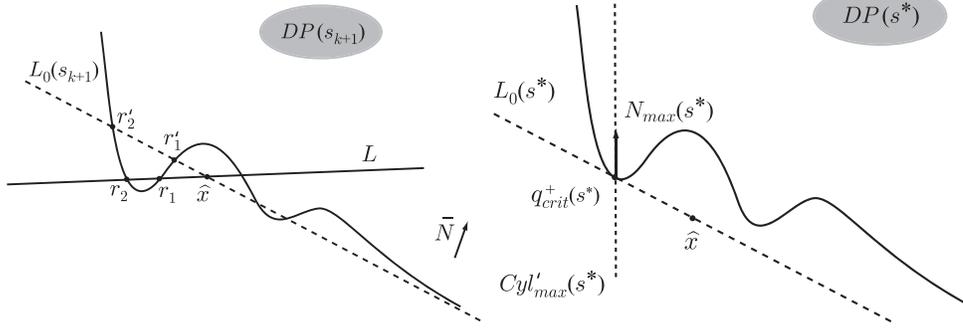
To prove (G2) observe that every point in the plane $x_3 = 1$ corresponds to a plane through x . Since the trajectory is complete, each plane through x intersects $C_{PI}(x)$ at least at one point, say $y(s_0)$. Then the B_{s_0} -line is tangent to the T -curve and passes through the chosen point in the plane $x_3 = 1$. The proof of (G3) is analogous to the proof of (G2). \square

APPENDIX C. PROOF OF PROPOSITIONS 4.5 AND 4.6

The proofs of propositions 4.5 and 4.6 are based on a number of auxiliary results. As usual, the IPs of Π with $C_{PI}(x)$ are denoted by $s_1 < s_2 < \dots < s_{2k+1}$.

Proposition C.1. *Consider $x \in U$ and $\Pi \in \Pi_{pr}(x)$. Then*

- (L1) $\bar{N} \cdot N_{max}(s_{k+1}) > 0$ if there is an IP in $IP(x, \Pi)$ to the right of $I_{PI}(x)$, and $\bar{N} \cdot N_{min}(s_{k+1}) < 0$ if there is an IP in $IP(x, \Pi)$ to the left of $I_{PI}(x)$;
- (L2) \hat{x} is below Γ_+ on $DP_+(s)$ and above Γ_- on $DP_-(s)$ for all $s \in I_{PI}(x)$;
- (L3) $\tau(l) \cdot \bar{N} > 0$ and $\tau(r) \cdot \bar{N} > 0$, where l and r are the smallest and largest IPs from $IP(x, \Pi)$, respectively;
- (L4) $IP(x, \Pi)$ contains an even number of IPs above $I_{PI}(x)$ and below $I_{PI}(x)$.


 FIGURE 24. Location of \hat{x} on $DP_+(s_{k+1})$

Proof. Recall that the normal \bar{N} to Π is chosen canonically (cf. (4.2)), so we shall refer to a point P as being *above* Π if $(P - y(s_1)) \cdot \bar{N} > 0$ and *below* Π if $(P - y(s_1)) \cdot \bar{N} < 0$.

We start by proving (L1). Suppose there is an IP $r \in IP(x, \Pi)$, which is located to the right of $IP_I(x)$, i.e. $r > s_t$. By construction, $r < q_{max}(s_{k+1})$. Thus $\bar{N} \cdot N_{max}(s_{k+1}) > 0$, because the points s_{k+1}, s_{2k+1}, r belong to $C(s_{k+1}, s_{2k+1}, r)$ and $r < q_{max}(s_{k+1})$. The argument here is exactly the same as the one used in establishing (4.2). So $y(q_{max}(s_{k+1}))$ is above Π . In the same way one shows that if there are IPs to the left of $IP(x, \Pi)$, then $y(q_{min}(s_{k+1}))$ is below Π .

We now prove (L2). To show that \hat{x} is below Γ_+ on $DP_+(s)$ when $s = s_b + \epsilon$, $0 < \epsilon \ll 1$ and \hat{x} is above Γ_- on $DP_-(s)$ when $s = s_t - \epsilon$, we note the following. Since x belongs to the line segment $H(s_b, s_t)$, \hat{x} always belongs to the projection of $H(s_b, s_t)$ onto $DP_{\pm}(s)$. When $s = s_b + \epsilon$ (resp. $s = s_t - \epsilon$), $\hat{H}(s_b, s_t)$ is a line segment on $DP_{\pm}(s)$, but when $s \rightarrow s_b$ (resp. $s \rightarrow s_t$), $\hat{H}(s_b, s_t)$ becomes a ray parallel to L_0 emanating from $\hat{y}(s_t)$ (resp. $\hat{y}(s_b)$) in the direction opposite to $\tau(s_b)$ (resp. $\tau(s_t)$). Since $s_t > q_{crit}^+(s_b)$ due to the assumption $x \in U$, this ray does not intersect Γ_+ . Otherwise we get a (local) critical chord $H(s_b, q)$, where $s_b < s_t < q$.

Let us discuss the last assertion in more detail. If there are additional IPs, let q_1 be the first IP above s_t . Suppose, for example, that the segment of Γ_+ located between s_t and q_1 is below the ray. Moving the ray down parallel to itself (and, of course, parallel to L_0), we find a line parallel to L_0 and tangent to Γ_+ at some point q , $s_t \leq q \leq q_1$. By construction, $H(s_b, q)$ is a critical chord, which gives the contradiction at the end of the previous paragraph. If the segment of Γ_+ is above the ray between s_t and q_1 , we get an analogous contradiction by moving the ray up.

Since \hat{x} on $DP_{\pm}(s)$ changes continuously, we conclude that when $s = s_b + \epsilon$, \hat{x} is below Γ_+ . In the same way one shows that \hat{x} is above Γ_- on $DP(s)$ when $s = s_t - \epsilon$. Next, we note that if \hat{x} belongs to Γ_{\pm} for some $s \in (s_b, s_t)$, then $y(s)$ is an endpoint of another PI line of x , which is impossible by the definition of U . Hence, \hat{x} is always below Γ_+ on $DP_+(s)$ and above Γ_- on $DP_-(s)$ for $s \in (s_b, s_t)$. It is necessary to define what below and above means when applied to Γ_{\pm} and \hat{x} on $DP_{\pm}(s)$. It follows from proposition 2.5 that for each s and for each $x \in Cyl'_{max}(s)$ (resp. $Cyl'_{min}(s)$) there exists a unique $t \in (s, q_{max}(s))$ (resp. $t \in (q_{min}(s), s)$) such that $[x - y(s), y(t) - y(s), N_{max}(s)] = 0$ (resp. $[x - y(s), y(t) - y(s), N_{min}(s)] = 0$), i.e. the corresponding line on $DP_+(s)$ (resp. $DP_-(s)$) intersects Γ_+ (resp. Γ_-) only once. This allows us to use the term above Γ_+ on $DP_+(s)$ (resp. below Γ_- on $DP_-(s)$) in the sense of using the direction $N_{max}(s)$ (resp. $-N_{min}(s)$) as the vertical axis.

Now, taking into account (L1), the locations of \hat{x} on $DP_{\pm}(s_{k+1})$, and the fact that $y(q_{max}(s_{k+1}))$ is above Π (due to (L1)), one sees that $\tau(l) \cdot \bar{N} > 0$ and $\tau(r) \cdot \bar{N} > 0$, which proves (L3).

We finally prove (L4). Π is a proper plane, so $y(s_t)$ is above Π . By (4.2), $N_{max}(s_1) \cdot \bar{N} > 0$. Since $y(s_1) \in \Pi$, $y(q_{max}(s_1))$ is also above Π . Hence, there is an even number of IPs of Π with $C(s_t, q_{max}(s_1))$. Similarly, one shows that there is an even number of IPs with $C(q_{min}(s_{2k+1}), s_b)$. Using that $\hat{y}(s_t)$ and $\hat{y}(q_{max}(s_{k+1}))$ are above L and $\hat{x} \in L$ is below Γ_+ on $DP_+(s_{k+1})$, we immediately get that there can only be an even number of IPs in $\Pi \cap C(s_t, q_{max}(s_1))$ to the left of \hat{x}_1 . Since the total number of IP in $\Pi \cap C(s_t, q_{max}(s_1))$ is also even, the assertion is proven. The assertions about the IPs below $IP_I(x)$ is proven analogously. \square

Proof of proposition 4.5. We will only prove the statement about r_p 's, the statement about l_p 's follows analogously. We will show that for all p , $1 \leq p \leq n$, $[x - y(s_{k+1}), y(r_p) - y(s_{k+1}), \bar{N}] < 0$, i.e. all IPs $y(r_p)$ are located on the right of the line through x and $y(s_{k+1})$ on Π (the orientation of Π is defined by the canonical choice of a normal vector). In other words, we need to show that all IPs r_1, \dots, r_n are located to the right of \hat{x} on $DP_+(s_{k+1})$. We denote the PI line of x by $H(s_b, s_t)$.

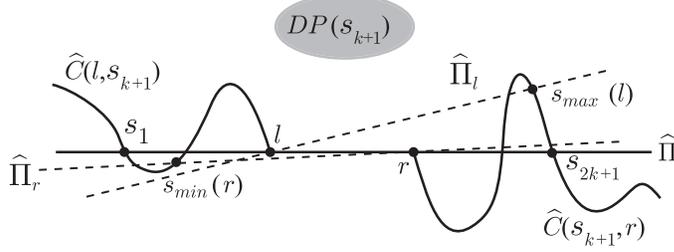
Suppose, on the contrary, that some of the r_p 's are to the left of \hat{x} on $DP_+(s_{k+1})$. Since $y(q_{max}(s_{k+1}))$ is above Π by (L1) and \hat{x} is below Γ_+ on $DP_+(s_{k+1})$ by (L2), there is an even number of such r_p 's (see Fig. 24, left panel). Denote the plane through x and touching C at $y(s)$ by $\Pi_0(s)$ and define $L_0(s) := \Pi_0(s) \cap DP_+(s)$ (note that s is variable). Define also $L := \Pi \cap DP_+(s_{k+1})$ and note that L has a greater slope than $L_0(s_{k+1})$ since s_{k+1} is an odd-indexed IP of Π with $C_{PI}(x)$. Hence, $\Pi_0(s_{k+1})$ also has at least two IPs with $C(s_{k+1}, q_{max}(s_{k+1}))$ to the left of \hat{x} on $DP_+(s_{k+1})$ (see Fig. 24, left panel). Indeed, $\Pi_0(s_{k+1})$ is obtained by rotating Π around $x - y(s_{k+1})$ in the clockwise direction, and since intermediate planes have at least two IPs with $C(s_{k+1}, q_{max}(s_{k+1}))$, we conclude that $\Pi_0(s_{k+1})$ has at least two IPs with $C(s_{k+1}, q_{max}(s_{k+1}))$ to the left of \hat{x} on $DP_+(s_{k+1})$ (this follows immediately from the shape of $\hat{C}(s_{k+1}, q_{max}(s_{k+1}))$, see proposition 2.5). Having established this property of $\Pi_0(s)$ for $s = s_{k+1}$, we consider $\Pi_0(s)$ and $L_0(s)$ on $DP_+(s)$ and decrease s from s_{k+1} towards s_b . First note that $s_t > q_{crit}^+(s_b)$ due to the fact that $x \in U$. This implies that $\Pi_0(s_b)$ has no IPs with $C(s_b, q_{max}(s_b))$ located to the left of \hat{x} on $DP_+(s_b)$ (cf. the proof of (L2) in proposition C.1). Hence, the last two IPs of $\Pi_0(s)$ with $C(s, q_{max}(s))$ to the left of \hat{x} on $DP_+(s)$ will collide at $q_{crit}^+(s)$ for some $s = s^* \in [s_b, s_t]$ (see Fig. 24, right panel). However, $x \in U$ implies that $x \in Cy'_{max}(t)$, $t \in [s_b, s_t]$. This gives the desired contradiction since \hat{x} is to the right of $y(q_{crit}^+(s^*))$ on $DP_+(s^*)$ and hence x does not belong to $Cy'_{max}(s^*)$. \square

Next we show that properties P1-P4 guarantee that IPs from the set $IP(x, \Pi)$ are not too far apart, i.e. they all belong to a turn of C . The first step towards this goal is to show that the ‘complicated’ part of $C(q_{min}(s), q_{max}(s))$ can always be projected onto a ‘detector plane’ tangent to C at $y(s)$ (see the discussion around properties P4(1) and P4(2) in section 2). The following proposition makes this statement precise:

Corollary C.2. *Consider $x \in U$ and $\Pi \in \Pi_{pr}(x)$. If $IP(x, \Pi)$ contains IPs both below $s_b(x)$ and above $s_t(x)$, then all IPs from $IP(x, \Pi)$ belong to the interval $(q_-(s), q_+(s))$.*

Proof. By (L1), $y(q_{max}(s_{k+1}))$ is above Π . Consider $DP_+(s_{k+1})$. If $q_-(s) = q_{max}(s)$, then also $q_-(s) = q_{min}(s)$ and the desired assertion follows from the definition of $IP(x, \Pi)$. Consider now the other case. We argue by contradiction. Suppose there is an IP $r \in IP(x, \Pi)$ above q_- , i.e. $r \in [q_-(s), q_{max}(s)]$. By construction, $\hat{y}(r)$ is to the right of \hat{x} . The assumption $x \in U$ implies that \hat{x} is below L_{max}^+ . Thus the slope of $DP_+(s) \cap \Pi$ is greater than the slope of L_0 , and there are no other IPs with $C(s, q_{max}(s))$ due to property P4(2) and the convexity of $\hat{C}(q_-(s), q_{max}(s))$. Consequently, r is the only one IP in $IP(x, \Pi)$ above s_t , which contradicts (L3). The case of IPs below s_b is analogous. \square

We now prove that IPs from $IP(x, \Pi)$ always belong to one turn of the curve.


 FIGURE 25. Illustration of the planes Π_l and Π_r

Proposition C.3. *Choose $x \in U$ and a plane $\Pi \in \Pi_{pr}(x)$. Denote the smallest IP from $IP(x, \Pi)$ by l and the largest IP by r . Then $r < q_{max}(l)$.*

Proof. Suppose r and l are separated by at least one turn of C . If $q_{max}(l) = r$, then $N_{max}(l)$ is parallel to Π , which contradicts (4.2) because $l \in (q_{min}(s_{2k+1}), s_1)$. Thus we assume $r > q_{max}(l)$. Consider the plane Π_l through $y(s_{k+1})$ and $y(l)$ and parallel to $N_{max}(l)$. Since $q_{max}(l) < r$, Π_l intersects $C(s_{k+1}, r)$ at least at one point $y(q_{max}(l))$, which is necessarily above Π by (4.2). It also follows from proposition 2.5 that Π_l does not intersect $C(l, s_{k+1})$.

Introduce the plane Π_r through $y(s_{k+1})$ and $y(r)$ and parallel to $N_{min}(r)$. Performing a similar analysis as above, we conclude that Π_r intersects $C(l, s_{k+1})$ at $y(q_{min}(r))$, $y(q_{min}(r))$ is below Π , and Π_r does not intersect $C(s_{k+1}, r)$.

Now one notes that by corollary C.2, $r < q_-(s_{k+1})$ and $l > q_-(s_{k+1})$. Since our proof will use only the curve segment $C(q_-(s_{k+1}), q_-(s_{k+1}))$, we can find a convenient detector plane onto which the segment projects well. It follows from properties P4(1) and P4(2) that $C(q_-(s_{k+1}), q_-(s_{k+1}))$ stays on one side of any plane, which contains $y(s_{k+1})$ and is parallel to $\tau(s_{k+1})$ and $(1 - \alpha)N_{max}(s_{k+1}) - \alpha N_{min}(s_{k+1})$, $0 \leq \alpha \leq 1$. Hence we define $DP(s_{k+1})$ as a plane which is parallel to $\tau(s_{k+1})$ and $(1 - \alpha)N_{max}(s_{k+1}) - \alpha N_{min}(s_{k+1})$ for some $0 \leq \alpha \leq 1$ and does not contain $y(s_{k+1})$. For convenience, we assume that U is between $DP(s_{k+1})$ and $y(s_{k+1})$.

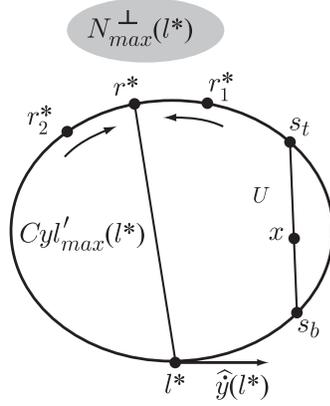
In the following proof we project $C(q_-(s_{k+1}), q_-(s_{k+1}))$ from $y(s_{k+1})$ onto $DP(s_{k+1})$. Define $\hat{\Pi}_l := DP(s_{k+1}) \cap \Pi_l$, $\hat{\Pi}_r := DP(s_{k+1}) \cap \Pi_r$. One sees that the lines $\hat{\Pi}_l$ and $\hat{\Pi}_r$ need to satisfy the following properties:

- (M1) $\hat{\Pi}_l$ contains l , does not intersect $\hat{C}(l, s_{k+1})$, and intersects $\hat{C}(s_{k+1}, r)$ above $\hat{\Pi}$;
- (M2) $\hat{\Pi}_r$ contains r , does not intersect $\hat{C}(s_{k+1}, r)$, and intersects $\hat{C}(l, s_{k+1})$ below $\hat{\Pi}$.

Additionally, $s_1 \in \hat{C}(l, s_{k+1}) \cap \hat{\Pi}$ and $s_{2k+1} \in \hat{C}(s_{k+1}, r) \cap \hat{\Pi}$. One sees that these properties cannot be satisfied simultaneously (see Fig. 25) unless Π_l , Π_r and Π are identical, which would contradict (4.2). \square

Proof of proposition 4.6. We prove this proposition by contradiction. Suppose for some $x \in U$ there exists a plane $\Pi \in \Pi_{pr}(x)$ such that $IP(x, \Pi)$ contains IPs both to the left and to the right of $I_{PI}(x)$. Using (L3), denote the two smallest IPs from $IP(x, \Pi)$ by $l_1 < l_2$, and the two greatest – by $r_1 < r_2$. By the assumption of the proposition, $s_b(x) > l_2$ and $s_t(x) < r_1$.

We first rotate Π around $\beta(s_{k+1}, x)$, i.e. around the line joining the middle IP s_{k+1} with x . By the definition of $IP(x, \Pi)$, l_1, l_2 are in the left halfplane of Π relative to this line, while r_1, r_2 are in the right halfplane. By (L2), the direction of rotation can be chosen so that l_1 increases and r_2 decreases. Consequently, l_2 decreases (and stays smaller than s_b), r_1 increases (and stays greater than s_t), and both $l_2 - l_1, r_2 - r_1$ decrease. We can continue to rotate until one of the pairs collapses into a single point (i.e. a tangency is achieved there) or a tangency occurs earlier at one of the

FIGURE 26. Rotation around $\tau(l^*)$

four IPs. Suppose, for example, that l^* is the point where the tangency is achieved, and r_1^*, r_2^* are the corresponding positions of r_1, r_2 after the rotation. Thus the plane Π^* through $l^*, r_1^*,$ and r_2^* is tangent to C at l^* . By construction, $l^* < r_1^* < r_2^* < q_{max}(l^*)$. Now rotate Π^* around $\tau(l^*)$. Because of the convexity of $\hat{C}(l^*, q_{max}(l^*))$ (see Fig. 26), the direction of rotation can be chosen to make r_1^* increase and r_2^* decrease. Again, we either get a tangency when r_1^* and r_2^* collapse into a single point or earlier at one of the two IPs. Let r^* be the resulting point of tangency. We have obtained a critical chord $H(l^*, r^*)$ with $l^* < s_b < s_t < r^*$. Hence, $x \notin Cy'_{max}(l^*)$, which contradicts the definition of U . \square

APPENDIX D. LIST OF MAIN SYMBOLS

Symbol	Definition
$\tau(s)$	The tangent vector to the curve C . It is defined by linear interpolation at points where the first derivative is discontinuous (see (A1)–(A3)).
$\Pi_{osc}(s)$	The osculating plane of C at $y(s)$, i.e. the plane through $y(s)$ containing $\tau(s)$ and $\dot{\tau}(s)$.
$H(s_0, s_1)$	The line segment with endpoints $y(s_0)$ and $y(s_1)$ ($s_0 < s_1$).
$C(s_0, s_1)$	The section of C with endpoints $y(s_0)$ and $y(s_1)$ ($s_0 < s_1$).
$Q(s_0, s_1)$	The function $Q(s_0, s_1) := [y(s_0) - y(s_1), \tau(s_1), \tau(s_0)]$ that is related to local uniqueness of PI lines.
$q_{max}(s)$ (resp. $q_{min}(s)$)	Smallest $s' > s$ (resp. largest $s' < s$) such that $H(s, s')$ is a global critical chord.
$N_{max}(s)$ (resp. $N_{min}(s)$)	Vector along the corresponding maximal PI line: $N_{max}(s) = y(q_{max}(s)) - y(s)$ (resp. $N_{min}(s) = y(q_{min}(s)) - y(s)$)
$I_{PI}(x) = [s_b, s_t]$	The PI-parametric interval of x , i.e. the interval $[s_b, s_t]$, where $H(s_b, s_t)$ is the PI-line of x .
$C_{PI}(x)$	The PI-parametric segment of C , i.e. $C(s_b, s_t)$.
$L_{PI}(x)$	The PI line of x , i.e. $H(s_b, s_t)$.
$\Pi_{max}(s)$ (resp. $\Pi_{min}(s)$)	The plane determined by the critical chord $H(s, q_{max}(s))$ (resp. $H(q_{min}(s), s)$), i.e. the plane containing $H(s, q_{max}(s))$ (resp. $H(q_{min}(s), s)$) and tangent to C at both endpoints of the chord.

Symbol	Definition
$DP_+(s)$ (resp. $DP_-(s)$)	Upper (resp. lower) ‘detector plane’, i.e. a plane parallel to $\Pi_{max}(s)$ (resp. $\Pi_{min}(s)$) and not passing through $y(s)$. The planes are chosen so that $L_0 = DP_+(s) \cap DP_-(s) \subseteq \Pi_{osc}(s)$.
Γ_+ (resp. Γ_-)	Stereographic projection of $C(s, q_{max}(s))$ on $DP_+(s)$ (resp. $C(q_{min}(s), s)$ on $DP_-(s)$) from $y(s)$.
\hat{C}	<ul style="list-style-type: none"> • Orthogonal projection of $C(q_{min}(s), q_{max}(s))$ onto the plane $\tau^\perp(s)$ in the main sections of the paper; • Orthogonal projection of $C(s, q_{max}(s))$ (resp. $C(q_{min}(s), s)$) onto the plane $N_{max}^\perp(s)$ (resp. $N_{min}^\perp(s)$) in appendix A.
$\hat{N}_{max}(s)$ (resp. $\hat{N}_{min}(s)$)	Projection of $N_{max}(s)$ (resp. $N_{min}(s)$) onto the plane $\tau^\perp(s)$.
L_{min}^+, L_{max}^-	Lines $L_{min}^+ = DP_+(s) \cap \Pi_{min}(s)$ and $L_{max}^- = DP_-(s) \cap \Pi_{max}(s)$ on the ‘detector planes’.
L_0	The intersection of $\Pi_{osc}(s)$ and $DP_+(s)$ or, equivalently, $DP_-(s)$
$q_+(s)$	The smallest $s < s' < q_{max}(s)$ such that $y(s') \in \Pi_{min}(s)$. Equal to $q_{max}(s)$ if there is no such s' .
$q_-(s)$	The largest $q_{min}(s) < s' < s$ such that $y(s') \in \Pi_{max}(s)$. Equal to $q_{min}(s)$ if there is no such s' .
$Cyl_{max}(s)$ (resp. $Cyl_{min}(s)$)	Infinite cylinder with base $\hat{C}(s, q_{max}(s))$ and axis $N_{max}(s)$ (resp. with base $\hat{C}(q_{min}(s), s)$ and axis $N_{min}(s)$).
$q_{crit}^+(s)$	The largest $s', s < s' < q_{max}(s)$ such that $H(s, s')$ is a local critical chord; equal to s if there is no such s' .
$q_{crit}^-(s)$	The smallest $s', q_{min}(s) < s' < s$ such that $H(s', s)$ is a local critical chord; equal to s if there is no such s' .
$Cyl'_{max}(s)$ (resp. $Cyl'_{min}(s)$)	The part of the cylinder $Cyl_{max}(s)$ (resp. $Cyl_{min}(s)$) that is cut off from it by the plane containing $N_{max}(s)$ (resp. $N_{min}(s)$) and passing through $y(q_{crit}^+(s))$ (resp. $y(q_{crit}^-(s))$) and $y(s)$.
$\Omega_{crit}(x)$	The set of planes through x that are tangent to $C_{PI}(x)$, or contain $L_{PI}(x)$, or contain a point where $C_{PI}(x)$ is not smooth.
$\Pi_*(x)$	The set of planes through x which are not in $\Omega_{crit}(x)$ and have three or more IPs with $C_{PI}(x)$.
T -curve of x	The curve on S^2 defined by $u(s) = \frac{(y(s)-x) \times \tau(s)}{ (y(s)-x) \times \tau(s) }$, $s \in I_{PI}(x)$.
A -curve of x	The great circle defined by $(y(s_t) - y(s_b))^\perp$.
B_s -curve	The great circle defined by $(y(s) - x)^\perp$.
$\Pi_{pr}(x)$	The set of <i>proper</i> planes (see definition 3.5).
$\Pi_{im}(x)$	The set of <i>improper</i> planes (see definition 3.5).
$IP(x, \Pi)$	The subset of $C \cap \Pi$ that includes all IPs located inside $I_{PI}(x)$ and possibly some extra IPs located outside $I_{PI}(x)$ (see section 4.2).
$\beta(s, x)$	Unit vector pointing from the current source position $y(s)$ to x .
$\alpha(\theta)$	Unit vector in the plane $\beta(s, x)^\perp$, parameterized by polar angle θ ; $\theta = 0$ corresponds to the plane containing the PI line, small $\theta < 0$ correspond to proper planes, small $\theta > 0$ correspond to improper planes.
$\Pi(\theta)$	Plane through $y(s)$ and parallel to $\beta(s, x)$ and $\alpha(\theta)$.

Symbol	Definition
$n(s, x, \alpha)$	Weight assigned to the IP $y(s) \in \Pi \cap C_{PI}(x)$, where Π is the plane through x parallel to $\beta(s, x)$ and α .
$q(\Pi)$	Function defined on the set of proper planes that controls the assignment of weights to IPs.
\mathcal{L}_1	Family of planes through $y(s)$ that contain $\tau(s)$.
\mathcal{L}_2	Family of planes through $y(s)$ that are tangent to $C_{PI}(x)$ for some $x \in U$.
\mathcal{L}_3	Family of planes through $y(s)$ that satisfy $q(\Pi) = 0$ for some $x \in U$ (the dependence of the definition of q on x is implicit, see section 4.3; note that the dependence preserves the FBP structure of the algorithm).
$L(t)$	Line tangent to C at $y(t)$.
$\hat{L}(t)$	Orthogonal projection of $L(t)$ onto the plane $N_{max}(s)^\perp$. It is used when $t \in (s, q_{max}(s))$.
$s_1 < s_2 < \dots < s_{2k+1}$	IPs of a plane $\Pi \in \Pi_{pr}(x)$ with $C_{PI}(x)$.
$l_1 < \dots < l_m$	IPs of $\Pi \in \Pi_{pr}(x)$ with the interior of $C(q_{min}(s_{2k+1}), s_b(x))$.
$r_1 < \dots < r_n$	IPs of $\Pi \in \Pi_{pr}(x)$ with the interior of $C(s_t, q_{max}(s_1))$.
$DP(s)$	A plane which is parallel to $\tau(s)$ and $(1 - \alpha)N_{max}(s) - \alpha N_{min}(s)$ for some $0 \leq \alpha \leq 1$ and does not contain $y(s)$. Note that only the stereographic projection of $C(q_-(s), q_+(s))$ on $DP(s)$ is guaranteed to be a piecewise smooth continuous curve. For example, the projection of $C(s, q_{max}(s))$ or $C(q_{min}(s), s)$ on $DP(s)$ need not be a continuous curve.

REFERENCES

- [BKP05] C. Bontus, T. Köhler, and R. Proksa, *EnPiT: filtered back-projection algorithm for helical CT using an n -Pi acquisition*, IEEE Transactions on Medical Imaging **24** (2005), 977–986.
- [CZLN06] G.-H. Chen, T.-L. Zhuang, S. Leng, and B. E. Nett, *Shift-invariant and mathematically exact cone-beam FBP reconstruction using a factorized weighting function*, IEEE Transactions on Medical Imaging (2006), submitted.
- [Kat02] A. Katsevich, *Analysis of an exact inversion algorithm for spiral cone-beam CT*, Physics in Medicine and Biology **47** (2002), 2583–2598.
- [Kat03] ———, *A general scheme for constructing inversion algorithms for cone beam CT*, International Journal of Mathematics and Mathematical Sciences **21** (2003), 1305–1321.
- [Kat04a] ———, *Image reconstruction for the circle and line trajectory*, Physics in Medicine and Biology **49** (2004), 5059–5072.
- [Kat04b] ———, *An improved exact filtered backprojection algorithm for spiral computed tomography*, Advances in Applied Mathematics **32** (2004), 681–697.
- [Kat04c] ———, *On two versions of a 3π algorithm for spiral CT*, Physics in Medicine and Biology **49** (2004), 2129–2143.
- [Kat05] ———, *Image reconstruction for the circle and arc trajectory*, Physics in Medicine and Biology **50** (2005), 2249–2265.
- [Kat06] ———, *$3PI$ algorithms for helical computer tomography*, Advances in Applied Mathematics **21** (2006), 213–250.
- [KBH04] A. Katsevich, Samit Basu, and Jiang Hsieh, *Exact filtered backprojection reconstruction for dynamic pitch helical cone beam computed tomography*, Physics in Medicine and Biology **49** (2004), 3089–3103.
- [KBK06] Th. Köhler, C. Bontus, and P. Koken, *The Radon-split method for helical cone-beam CT and its application to nongated reconstruction*, IEEE Transactions on Medical Imaging **25** (2006), 882–897.
- [KK06] M. Kapralov and A. Katsevich, *A one-PI algorithm for helical trajectories that violate the convexity condition*, Inverse Problems **22** (2006), 2123–2143.

- [KK07] A. Katsevich and M. Kapralov, *Filtered backprojection inversion of the cone beam transform for a general class of curves*, SIAM Journal on Applied Mathematics (2007), to appear.
- [PN05] J. D. Pack and F. Noo, *Cone-beam reconstruction using 1D filtering along the projection of M-lines*, Inverse Problems **21** (2005), 1105–1120.
- [PNC05] J. D. Pack, F. Noo, and R. Clackdoyle, *Cone-beam reconstruction using the backprojection of locally filtered projections*, IEEE Transactions on Medical Imaging **24** (2005), 1–16.
- [SZP05] E.Y. Sidky, Y. Zou, and X. Pan, *Minimum data image reconstruction algorithms with shift-invariant filtering for helical, cone-beam CT*, Physics in Medicine and Biology **50** (2005), 1643–1657.
- [YLKK06] H. Yang, M. Li, K. Koizumi, and H. Kudo, *Exact cone beam reconstruction for a saddle trajectory*, Physics in Medicine and Biology **51** (2006), 1157–1172.
- [YZYW05a] Y. Ye, S. Zhao, H. Yu, and G. Wang, *A general exact reconstruction for cone-beam CT via backprojection-filtration*, IEEE Transactions on Medical Imaging **24** (2005), 1190–1198.
- [YZYW05b] H. Yu, S. Zhao, Y. Ye, and G. Wang, *Exact BPF and FBP algorithms for nonstandard saddle curves*, Medical Physics **32** (2005), 3305–3312.
- [ZLNC04] T. Zhuang, S. Leng, B. E. Nett, and G. Chen, *Fan-beam and cone-beam image reconstruction via filtering the backprojection image of differentiated projection data*, Physics in Medicine and Biology **49** (2004), 1643–1657.
- [ZPXW05] Y. Zou, X. Pan, D. Xia, and G. Wang, *PI-line-based image reconstruction in helical cone-beam computed tomography with a variable pitch*, Medical Physics **32** (2005), 2639–2648.