On Finite Rotations and the Noncommutativity Rate Vector

L.P. Candy and J. Lasenby

***Abstract*—The orientation vector differential equation was first derived by John Bortz to improve the accuracy of strapdown inertial navigation attitude algorithms. These algorithms previously relied on the direct integration of the direction cosine matrix differential equation. Here a compact derivation of the Bortz equation using geometric algebra is presented. Aside from being as simple and direct as any derivation in the literature, this derivation is also entirely general in that it yields a form of the Bortz equation that is applicable in any dimension, not just the conventional 3D case. The derivation presented has the further advantage that it does not rely on multiple methods of representing rotations and is expressed in a single algebraic framework.**

In addition to the new derivation, the validity of the notion that it is the effect of the noncommutativity of finite rotations that necessitates the use of such an equation in strapdown inertial navigation systems (SDINS) is questioned, and alternative justification for using the Bortz equation is argued.

I. Introduction

In strapdown inertial navigation systems (SDINS) the angular velocity measured in a body frame is used to update a rotation generator that relates the orientation of the body frame to some reference frame. There is a problem with this method however in that the angular velocity measurements are made in the rotating body frame, and the plane, or axis, of rotation over a finite time interval is not constant with respect to the reference frame. This results in the direct integration of the direction cosine differential equation being an unsuitable method for tracking attitude, an effect that has traditionally been attributed to the noncommutative nature of finite rotations [3]–[7]. This interpretation is however misleading, since while it is true that finite rotations do not commute, the source of the error in this case lies in the behavior of the measurement frame and the nature of the measurement.

Bortz derived a method for accounting for this effect [3] by representing the actual finite rotation by an orientation vector $\phi(t)$ and then obtaining an expression for $\dot{\phi}(t)$. Bortz showed that $\dot{\phi}(t)$ has two components, $\omega(t)$ and $\dot{\sigma}(t)$, where $\omega(t)$ is the inertially measured rate vector, and $\dot{\sigma}(t)$ is what Bortz termed the *non-inertially measurable non-commutativity rate vector*. Bortz's derivation is somewhat lengthy and involved and a number of subsequent derivations have been published [4]–[6], [8].

This paper supports a contrasting interpretation to that commonly presented regarding the requirement for the Bortz equation (section II) which is in line with the work of Jekeli [2]. Next a brief introduction to geometric (Clifford) algebra is presented (section III). This serves to introduce the mathematical formalism used in section IV to present a concise derivation of the Bortz equation that is more general than those that have been presented previously [3]–[6], [8]. The derivation presented is as compact as Savage's derivation in [8] but where the dynamical equation in [8] is relatively complex to obtain [2], obtaining the dynamical equation in the formalism presented here is trivial [12].

Readers familiar with geometric algebra may wish to skip section III.

We note that this same problem of updating rotations also occurs in other fields, for example, mechanical finite element analysis where other forms of the Bortz equation and its derivation can be found [14].

Liam Candy is a researcher with the Council for Scientific and Industrial Research, South Africa, and a member of the Signal Processing group at Cambridge University Engineering Department.

Joan Lasenby is a lecturer in the Engineering Department at Cambridge University and a Fellow of Trinity College.

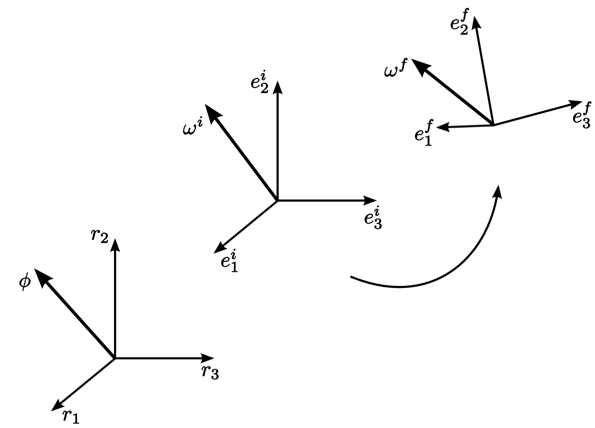

Fig. 1. A general rotation interval where the initial frame, e^i, rotates to the final frame, e^f, under the influence of ω, where ω^i and ω^f are the time-varying angular rate vectors in the e^i and e^f frames respectively. For any finite time interval, a rotation vector ϕ exists in the reference frame that will achieve the same rotation as ω acting over the interval.

II. Finite Rotations

The conventional method for tracking the orientation of the body frame relative to the reference frame[1] in SDINS applications is through the use of a rotation matrix C, or its quaternion counterpart, which will take a vector from the body frame to the reference frame. Before Bortz published his 1971 paper [3] the standard method for computing C from angular rate measurements was to integrate the equation[2]

$$\dot{C} = C[\omega\times] \tag{1}$$

where $[\omega\times]$ is the skew symmetric matrix form of the vector representing the angular velocity of the body frame with respect to the reference frame, expressed in the body frame, such that for any vector a, $[\omega\times]\,a = \omega \times a$. This method of updating C, however, leads to an attitude error that accumulates with time, an effect that has historically been attributed to the noncommutative nature of finite rotations. While it is demonstrable that finite rotations do not commute [1], an explicit link between this effect and the errors encountered in direct integration of equation 1 appears to be absent in the literature.

In a standard inertial navigation system, three orthonormal gyroscope measurements are combined into a single angular velocity vector which gives the angular rate of the body with respect to an inertial reference frame. The rotation generated by integrating this vector is a single rotation. In other words, there is no built-in ordering assumed in terms of the roll, pitch and yaw gyroscope measurements, they are handled simultaneously. Secondly, the series of rotations is sequential in time and so can only be meaningfully applied in a set order. Consequently, it is difficult to see how the non-commutative nature of finite rotations is applicable, as there is no set of rotations which may be applied 'out of order'.

[1] The reference frame may be any arbitrarily chosen inertial frame.

[2] Current algorithms rely on integrating one or another form of the Bortz equation to compute ϕ, and then use the fact that $C = f(\phi)$ to obtain C without integrating equation 1 directly.

The fundamental problem with integrating equation 1 directly is that gyroscopes (and SDINS algorithm implementations) have a finite bandwidth and their measurement frame is rotating. The differential equation derived by Bortz is no more 'analytically exact' than equation 1, but it achieves better results in discrete-time implementations because it explicitly accounts for the fact that the measurement frame is rotating via the $\dot{\sigma}$ term.

Consider a body rotating under the influence of the time-varying (measured in the body frame) angular rate ω shown in Fig. 1. Under the influence of this measured rate the body axes e^i rotate to e^f in some finite time interval. Equation 1 will update C by rotating it about the axis defined by ω fixed in the reference frame. This poses a difficulty however, since while ω is known in the body frame, and the initial orientation of the body frame with respect to the reference frame may be known, the orientation of the body frame with respect to the reference frame is changing with time by definition ($\omega \neq 0$). The result is an error which accumulates with every integration cycle, since the orientation of ω is not correct for the entire integration interval. (Clearly this error is not present for special case where the direction of ω is fixed - i.e. where there is no *coning* motion [2].)

It is always possible to express the actual rotation in the scenario above using a vector ϕ, where a rotation of $|\phi|$ about $\hat{\phi}$ will take the e^i frame to the e^f frame. Simply taking the derivative of ϕ and allowing for the fact that $\frac{d}{dt}\hat{\phi}$ is not necessarily 0 provides an equation for $\dot{\phi}$ that is dependent on the angular velocity about the vector ϕ as well as the angular velocity of the axis defined by ϕ in the reference frame. This is clear mathematically in equation 17, and all that is required to derive the Bortz equation is to solve for these two quantities in terms of the measured angular rate ω.

Many authors [1], [3], [5]–[7], including Bortz himself, have justified the use of the Bortz equation by claiming that it accounts for the noncommutative nature of finite rotations. Given the above, it would seem that labeling $\dot{\sigma}$ the *noncommutativity rate vector* or even describing it as *non-inertially measurable* is somewhat misleading, as not only is the complete angular rate of the body inertially measurable, but the need to compensate for body rotation with $\dot{\sigma}$ does not arise from the fact that finite rotations do not commute. Rather the $\dot{\sigma}$ term compensates for the motion of the ω axis – this is discussed further in section IV.

Jekeli [2] attributes the requirement for the Bortz equation as arising from the fact that angular rate measurements are being made in the rotating body frame and not as a result of the fact that finite rotations do not commute. He still employs the term *commutativity error* since it is possible to show that in a quaternionic formulation of the attitude equations, in the case that $A\dot{A} = \dot{A}A$, where A is a 4×4 skew symmetric matrix of the angular rates, we have no coning which results in the elimination of the third order term in the algorithm error. Stating that A commutes with its derivative is, as Jekeli notes, simply a statement of the fact that $\frac{d}{dt}\hat{\omega} = 0$, i.e. the direction of ω does not change.

III. Geometric Algebra

This section provides a brief introduction to geometric algebra. Geometric algebra is a mathematical language that is rapidly finding new application in engineering, physics and computer graphics [10]–[13]. The introduction that follows will be limited to a 3-dimensional space with orthonormal basis $[e_1 e_2 e_3]$ and positive signature[3], however, the concepts introduced are applicable to spaces of any dimension and signature.

[3]The signature of a space is positive if all of its basis vectors have positive squares. An example of a space of mixed signature is Minkowski space-time, where the three spatial dimensions have positive signature but the fourth basis squares to -1, i.e. it has a negative signature.

We will begin our introduction to geometric algebra with a new product for vectors; the outer product. The outer product of two vectors, a and b, returns a *directed area element* $a \wedge b$, called a *bivector*. The bivector can be visualised by sweeping the vector b along the length of the vector a to form a directed parallelogram with area $|a|\,|b| \sin\theta$, where θ is the angle between the vectors. This idea of a directed area can be generalised to higher dimensions by continued 'wedging'; the bivector $a \wedge b$ and the vector c can be wedged to produce a directed volume, or trivector, $a \wedge b \wedge c$, and so on. As the outer product of a vector with itself is zero, elements of higher grade than the dimension of the space cannot exist. In the case of the 3-dimensional space defined above, the only trivector, or grade-3 element, up to scale, is $e_1 \wedge e_2 \wedge e_3$. This highest grade element in a space is known as the pseudoscalar and is denoted as I. Lastly, an important property of the outer product is that it is anticommutative, i.e. $a \wedge b = -b \wedge a$. Note that for vectors, $a \wedge b = I\,(a \times b)$.

We are now in a position to define the geometric product of vectors as

$$ab = a \cdot b + a \wedge b \tag{2}$$

where $a \cdot b$ is the familiar vector dot product which we will call the inner product. We regard the geometric product as a more fundamental product of vectors than the inner or outer products. Using equation 2 and the fact that the outer product is anticommutative we can redefine the inner and outer products in terms of the geometric product as follows:

$$a \cdot b = \frac{1}{2}\,(ab + ba) \tag{3}$$

$$a \wedge b = \frac{1}{2}\,(ab - ba) \tag{4}$$

Geometric algebra is a graded algebra, where the *grade* of an element is equal to its dimension. So a scalar within the algebra has grade-0, a vector has grade-1, a bivector has grade-2 and so on. As the outer product of several vectors is called a *blade*, the bivector $a \wedge b$ is an example of a blade of grade 2. The geometric product for a vector and any blade of grade p can be generalised as

$$aA_p = a \cdot A_p + a \wedge A_p \tag{5}$$

from which we can obtain

$$a \cdot A_p = \frac{1}{2}\,(aA_p - (-1)^p\,A_p a) \tag{6}$$

$$a \wedge A_p = \frac{1}{2}\,(aA_p + (-1)^p\,A_p a) \tag{7}$$

A *multivector* is a linear combination of elements of any grade. While at first it seems odd to think of the sum of scalars, vectors, directed areas and directed volumes, it is really no more odd than the sum of a real and imaginary number forming a complex number. In general, any product of two multivectors returns the sum of the products of every blade in the first multivector with every blade in the second. The grade of the inner and outer products of two blades is given by the difference and sum respectively of the grades of the blades involved. The inner product is therefore also known as the grade-lowering product and the outer product as the grade-raising product.

There are two other important operators to be introduced before continuing. The first is the *reversion* operator, which has the effect of reversing the order of vectors in a product and is denoted by a tilde. In other words if $c = ab = a \cdot b + a \wedge b$, then

$$\tilde{c} = ba = b \cdot a + b \wedge a = a \cdot b - a \wedge b \tag{8}$$

It is easy to show that in the 3-dimensional case the effect of reversion on a given multivector is to change the sign of the bivector and trivector components of the multivector.

The second operator to introduce is the *project* operator, denoted by $\langle M \rangle_x$. This expression will return the grade-x blades of the multivector M. The scalar component of M is denoted simply as $\langle M \rangle$.

Using equations 5, 6 and 7, it is possible to form any product in any vector space. It is possible to show, for example, that the product of two bivectors is given by[4]

$$\begin{aligned} AB &= A \cdot B + A \underline{\times} B + A \wedge B \\ &= \langle AB \rangle + \langle AB \rangle_2 + \langle AB \rangle_4 \end{aligned} \tag{9}$$

where $A \underline{\times} B = \frac{1}{2}(AB - BA)$. This product of bivectors reveals a general property of the inner and outer products: the grade of the inner product of two blades is the difference of the grades of the blades, while the grade of the outer product of two blades is the sum of the grades of the blades.

Geometric algebra allows for a very general method of representing rotations, this comes from the fact it offers a naturally superior quantity, when compared to vector algebra or matrix algebra [9], for representing rotation angles and rotational velocities: the bivector. By defining rotations in terms of a *plane of rotation* it is possible to avoid the situation that arises in vector algebra where an axis of rotation is undefined (for 2D) or ambiguous (for 4D and above). As a result we have a method of handling rotations and specifying rotation angles or angular velocities in n-dimensions.

It is easy to verify that all unit bivectors square to -1 in a space with positive signature, that is their geometric product is -1. This leads to a development analogous to the Euler representation of rotations in the complex plane; rotations in general space can be represented as exponentiated bivectors, termed *rotors*. A rotor that will produce a counter-clockwise rotation of α in the unit plane B is given by the expression

$$R = \mathrm{e}^{-\frac{B\alpha}{2}} = \cos\frac{\alpha}{2} - B\sin\frac{\alpha}{2} \tag{10}$$

The expression

$$a' = Ra\tilde{R} \tag{11}$$

will then rotate a in the plane B, through an angle α to give a'. Equation 11 will transform not only vectors, but any blade, while preserving the grade. The form of 11 should look familiar, and indeed, in the 3-dimensional case rotors can be shown to operate in exactly the same way as quaternions. A further property of rotors is that $R\tilde{R} = 1$, just as $qq^* = 1$ for their quaternion counterparts and $CC^T = 1$ for rotation matrices.

For the purposes of tracking the orientation of a rigid body with respect to some reference frame, the time dependent rotor $R(t)$ that will take vectors from the reference frame to the body frame can be obtained from [10]

$$\dot{R} = -\frac{1}{2}\Omega^r R \tag{12}$$

where Ω^r is the angular velocity of the rigid body with respect to the reference frame, expressed as a bivector in the reference frame (the bivector defines the plane of rotation and has a magnitude equal to the speed of rotation). It is worth noting the similarity between equation 12 and the quaternion counterpart of equation 1 [2].

[4]Clearly in a 3-dimensional case $A \wedge B = 0$, so equation 9 becomes $AB = A \cdot B + A \underline{\times} B = \langle AB \rangle + \langle AB \rangle_2$.

It is easy to show that if we express the angular velocity of the body with respect to the reference frame as a bivector in the body frame, Ω^b, then equation 12 becomes

$$\dot{R} = -\frac{1}{2}R\Omega^b \tag{13}$$

A second expression for $\dot{R}$ is obtainable via direct differentiation of equation 10

$$\dot{R} = -\frac{\dot{\alpha}}{2}\sin\frac{\alpha}{2} - \dot{B}\sin\frac{\alpha}{2} - B\frac{\dot{\alpha}}{2}\cos\frac{\alpha}{2} \tag{14}$$

IV. A Concise Derivation of the Bortz Equation

For the derivation that follows the superscript for Ω^b in equation 13 has been omitted, i.e. $\Omega^b \equiv \Omega$.

It is possible to specify any finite rotation in a unit plane, defined by the unit bivector B, and through an angle α using a rotation bivector;

$$\Phi = \alpha B \tag{15}$$

From this rotation bivector the corresponding rotor is easily obtained via,

$$R = \mathrm{e}^{-\frac{\Phi}{2}} = \mathrm{e}^{-\frac{\alpha B}{2}} \tag{16}$$

Differentiating equation 15 with respect to time gives,

$$\dot{\Phi} = \alpha\dot{B} + \dot{\alpha}B \tag{17}$$

The reference frame equation 17 shows that the angular rate that will result in the rotation Φ over some finite time interval consists of two components; one in the plane of rotation, $\dot{\alpha}B$, and a second component, $\alpha\dot{B}$, perpendicular to the plane of rotation. $\dot{B}$ is ‘perpendicular’ to B in the sense that $B \cdot \dot{B} = 0$; this can be seen from the fact that $|B^2| = 1$ implies $B\dot{B} + \dot{B}B = 0$ and therefore $B \cdot \dot{B} = 0$.

In order to complete the derivation, expressions for $\dot{B}$ and $\dot{\alpha}$ in terms of the measured angular rate Ω are all that is required. This can be achieved by simply equating the scalar and bivector components of equations 13 and 14.

Finding an expression for $\dot{\alpha}$: Equating the scalar components of 13 and 14 gives:

$$\langle R\Omega \rangle = \dot{\alpha}\sin\frac{\alpha}{2} \tag{18}$$

Noting that $\langle R\Omega \rangle = \Omega \cdot \langle R \rangle_2$, and substituting for the bivector component of R:

$$\dot{\alpha} = -\Omega \cdot B \tag{19}$$

Finding an expression for $\dot{B}$: Equating the bivector components of 13 and 14 gives:

$$\langle R\Omega \rangle_2 = 2\dot{B}\sin\frac{\alpha}{2} + \dot{\alpha}B\cos\frac{\alpha}{2} \tag{20}$$

Using the fact that $\langle R\Omega \rangle_2 = \langle R \rangle \Omega + \left\langle \langle R \rangle_2 \Omega \right\rangle_2$, and substituting for $\langle R \rangle$ and $\langle R \rangle_2$ from equation 10 and for $\dot{\alpha}$ from equation 19:

$$\begin{aligned} 2\dot{B}\sin\frac{\alpha}{2} &= \Omega\cos\frac{\alpha}{2} - \langle B\Omega \rangle_2 \sin\frac{\alpha}{2} - \dot{\alpha}B\cos\frac{\alpha}{2} \\ \dot{B} &= \frac{1}{2}\cot\frac{\alpha}{2}\left[\Omega + (\Omega \cdot B)B\right] - \frac{\langle B\Omega \rangle_2}{2} \end{aligned} \tag{21}$$

Completing the derivation: It is now possible to substitute the expressions for $\dot{B}$ and $\dot{\alpha}$ into the expression for $\dot{\Phi}$ in equation 17:

$$\dot{\Phi} = \Omega + \left(\frac{\alpha}{2}\cot\frac{\alpha}{2} - 1\right)\left[\Omega + (\Omega \cdot B)\, B\right] - \frac{\alpha}{2}\langle B\Omega\rangle_2 \tag{22}$$

Or, alternatively, by substituting for $B = \frac{\Phi}{|\Phi|}$ and $\alpha = |\Phi|$, it is possible to write equation 22 in terms of Φ only:

$$\dot{\Phi} = \Omega + \left(\frac{|\Phi|}{2}\cot\frac{|\Phi|}{2} - 1\right)\left[\Omega + \frac{(\Omega \cdot \Phi)\,\Phi}{|\Phi|^2}\right] - \frac{\langle \Phi\Omega\rangle_2}{2} \tag{23}$$

In the case where there is no coning motion the axis of rotation is stationary – i.e. its direction is fixed in space – which implies $\dot{B} = 0$ and B and Ω lie in the same plane, and in turn, from equations 17, 19 and 23, $\dot{\Phi} = \dot{\alpha}B = \Omega$.

Equation 23 is the final form for the Bortz equation expressed in geometric algebra. It is interesting to note that the starting point of the derivation, equations 17 and 13, are applicable to spaces of any dimension and at no point in the above derivation was the limitation of a 3-dimensional case assumed. While the work of Bar-Itzhack [9] extends Eulers theorem to n-dimensions, the Euler analogue in geometric algebra is given by equation 10, and is inherently general in n-dimensions. The result in equation 23 extends the Bortz equation to a form that is applicable to n-dimensions.

Equation 23 can be used directly in INS algorithms implemented using geometric algebra. Alternatively, the traditional Bortz equation can be obtained by substitution[5] of $\Omega = I\omega$, $\Phi = I\phi$, multiplication of both sides by the pseudoscalar, I, and simplification. For the sake of completeness this is carried out below, bearing in mind that $|\Phi| = |\phi|$ by definition;

$$\begin{aligned}
\dot{\phi} &= \omega + \left(\frac{|\phi|}{2}\cot\frac{|\phi|}{2} - 1\right)\left[\omega + \frac{(I\omega \cdot I\phi)\,\phi}{|\phi|^2}\right] - \frac{\langle I\phi I\omega\rangle_2 I}{2} \\
&= \omega + \left(\frac{|\phi|}{2}\cot\frac{|\phi|}{2} - 1\right)\left[\frac{|\phi|^2\,\omega - (\omega \cdot \phi)\,\phi}{|\phi|^2}\right] + \frac{\phi \times \omega}{2} \\
&= \omega + \frac{1}{2}\phi \times \omega + \frac{1}{|\phi|^2}\left(1 - \frac{|\phi|}{2}\cot\frac{|\phi|}{2}\right)\phi \times \phi \times \omega
\end{aligned} \tag{24}$$

as it is not hard to show that $\phi \times (\phi \times \omega) = -|\phi|^2\,\omega + (\omega \cdot \phi)\,\phi$. Equation 24 is the familiar form of the Bortz equation.

V. Conclusion

Geometric algebra is shown to be a useful framework in which to derive the Bortz equation. Aside from providing an extremely compact derivation, the form of the equation derived is entirely general. Regarding the requirement for the Bortz equation, the argument is made that it is a consequence of performing rate measurements in a rotating frame and not a consequence of the noncommutative nature of finite rotations.

References

[1] D Titterton and J Weston, *Strapdown Inertial Navigation Technology, 2nd Edition*, IEE, 2004.

[2] C Jekeli, *Inertial Navigation Systems with Geodetic Applications*, de Gruyter, 2001.

[3] J.E Bortz, "A New Mathematical Formulation for Strapdown Inertial Navigation", *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-7, no. 1, pp 61-66, January 1971.

[4] G.J Nazaroff, "The Orientation Vector Differential Equation", *AIAA Journal of Guidance, Control and Dynamics*, vol. 2, no. 4, pp 351-352, July-August 1979.

[5] Y.F Jiang and Y.P. Lin, "On the Rotation Vector Differential Equation", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 27, no. 1, pp 181-183, January 1991.

[6] M.B Ignani, "On the Orientation Vector Differential Equation in Strapdown Inertial Navigation Systems", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 4, pp 1076-1081, October 1994.

[7] L.E Goodman and A.R Robinson, "Effect of Finite Rotations on Gyroscopic Sensing Devices", *Journal of Applied Mechanics*, June 1958.

[8] M.D Shuster, "The Kinematic equation for the Rotation Vector", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 1, pp 263-267, January 1993.

[9] I.Y Bar-Itzhack, "Extension of Eulers Theorem to n-Dimensional Spaces", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 6, pp 903-909, November 1989.

[10] D Hestenes, "New Foundations for Classical Mechanics (Fundamental Theories of Physics)", Springer, 1999.

[11] D Hestenes and G Sobczyk, "Clifford Algebra to Geometric Calculus: A unified language for mathematics and physics", D. Reidel, Dortrecht, 1984.

[12] C.J.L Doran and A.N Lasenby, "Geometric Algebra for Physicists", Cambridge University Press, 2003.

[13] J Vince, "Geometric Algebra for Computer Graphics", Springer, 2008.

[14] F.A McRobie and J Lasenby, "Simo-Vu Quoc rods using Clifford algebra", *International Journal for Numerical Methods in Engineering*, vol. 45, no. 4, pp 377-398, 1999.

[5] The dual nature of bivectors and vectors can be established from equations 5, 6 and 7: $Ia = A$ and $-IA = a$, where the vector a has the same magnitude as the bivector A and has a direction normal to the plane defined by A -