دور المنشأ في تطورات عمارة القرن العشرين في العراق

الأستاذ عاطف السهيري	اسامة عبد المنعم
جامعة بغداد – كلية الهندسة	جامعة بغداد – كلية الهندسة

ABSTRACT

It may by said that there is a historical relationship between architecture and structure which was strongly affected by the technological developments. Therefore structure as a factor of the material base of architectural technology is very influencing factor in design activity, and in architectural configuration.

Many, other points of views and transformations that appeared casted their shadows upon the Iraqi architecture of 20th century, each period carried the entry of architectural and structural styles which caused a variation in effect, So the knowledge gap that the research focuses on the absence of the role of structure (on the practical level as a direct influencing factor) in the shaping of Iraqi architecture.

The subject matter is dealt with through an inductive research for the purpose of establishing basic concepts of the nature of structural effect upon styles and stresses the theoretical and practical integration in the developments of the 20th century architecture. The research reveals the large effect of structural system on architectural form according to the principle (cause- results). And the research concluded that:

- There are many influences factors (especially in structural technology) contributed to the shaping and styling of architecture in every period and place.
- There is a large effect upon architectural configuration (on practical and theoretical level) by structure in the various periods of Iraqi architecture, then concluded the levels of the nature of structural effect on architectural form, and their applications upon local architectural development.

الخـلاصــة:

يمكن القول بان هنالك علاقة تاريخية بين العمارة والمنشأ تأثرت كثيراً بالتطورات التكنولوجية خاصة في الجانب المنشئي منها، سواء على مستوى المواد الجديدة وخصائصها أو على مستوى وسائل الإنشاء. فالمنشأ كمرتكز مادي يشكل جانب أساسي من جوانب الفعالية التصميمية باعتباره الجزء الأكثر أهمية في التشكيل المعماري.

وقد ألقت الكثير من التوجهات والتحولات بظلالها على نتاج عمارة القرن العشرين في العراق حيث حملت كل حقبة دخول طرز معمارية ومنشئية الأمر الذي سبب تفاوتاً في وقع التأثير ومن ثم التحول إلى جزء من العمارة المحلية وأسلوبها التصميمي، وبالتالي فقد تركزت المشكلة البحثية في غياب التركيز حول دور المنشأ واثره على الشكل المعماري في عمارة القرن العشرين في العراق.

وعليه فقد تم تناول الموضوع من خلال بحث استقرائي تحليلي يهدف إلى وضع المفاهيم الأساسية لأنماط طبيعة التأثير المنشئي ووصفها وتأكيد الترابط الفكري والتطبيقي على مستوى العمارة المحلية. وبيان الأثر المباشر للمنظومة المنشئية في الشكل المعماري وفق علاقة

(مؤثر – نتيجة) باعتبار المنشأ الجزء الأكثر أهمية في التشكيل المعماري. وبعد مناقشة العديد من الأفكار ذات العلاقة بالموضوع واسقاطاتها في مراحل العمارة المحلية توصل البحث إلى مجموعة من النتائج أهمها:

- هنالك عدة مؤثرات ساهمت في تحديد صيغ وتوجهات كل عمارة في زمان ومكان خاصة ما تعلق منها بالتكنولوجيا المنشئية والتي امتازت بالديناميكية التطورية خلال الزمن.
- وقد افرز البحث التأثير الكبير الذي يفرضه المنشأ على المستويين الفكري والتطبيقي في التشكيل المعماري وفي مختلف مراحل تطور العمارة العراقية. حيث تم استخلاص عدة مستويات لطبيعة التأثير المنشئي والذي اتخذ بدوره عدة توجهات تبعاً لطبيعة العلاقة التأثيرية وتطبيقاتها على تطورات العمارة المحلية، والتي أثرت في نتاج كل مرحلة.

The research focuses on establishing the nature of the role of structure in the 20th century architecture in Iraq.

دور المنشأ في تطورات عمارة القرن العشرين في العراق :

- **المقدمة:**

تعّد العمارة نتاج لمراحل متتالية ذات حلقات مترابطة متكاملة تسعى لبناء الفكر المعماري.... حيث تتداخل عندها القيم الاجتماعية والاقتصادية والسياسية وحتى الأكاديمية للمجتمع مروراً بالقيم التكنولوجية والفنية على مختلف مستوياتها، فالعمارة على مختلف العصور تعد المرآة التي تنعكس عليها الأوضاع الاجتماعية والاقتصادية والتقنية للمجتمعات في كل عصر. حيث يظل تلازم الشكل المعماري بالتقنية والتطورات المنشئية الحديثة أحد الهواجس الأساسية التي يبحث عن مساحات جديدة لها كل مهتم بالعمارة، لأنه من غير المتوقع أن تظهر عمارة رائدة وذات قيمة دون أن تلازمها بالتقنية المتطورة التي تعبر عن آخر ما توصل إليه العلم في مجالات الصناعة البنائية والمنشئية .. وهو ما يبرز أهمية التطور التقني والعلم وتأثيره المباشر على الفعالية البنائية عموماً.. وقد شهد القرن العشرين ظهور الكثير من التحولات والأفكار التي امتازت بسرعة تحولاتها الأيديولوجية والتطبيقية في مجال التكنولوجيا البنائية عموماً والمنظومة المنشئية على وجه الخصوص مما أنتج العديد من التوجهات والتحولات المعمارية التي ألقت بظلالها على نتاج عمارة القرن العشرين في العراق.

وتمثلت المشكلة البحثية هنا في غياب التركيز حول طبيعة دور المنشأ على الشكل المعماري في عمارة القرن العشرين في العراق وقلة الدراسات المحلية حوله. مع غياب التطرق للأثر والمحددات المعمارية والشكلية المرتبطة بالمنشأ وخاصة ما يتعلق بالجانب الحدسي.

وعليه يهدف البحث في المحاولة في إبراز طبيعة وخصائص الدور المنشئي ووصفها وتأكيد الترابط الفكري والتطبيقي على مستوى العمارة المحلية خلال سنوات القرن العشرين وفق تصنيف مرحلي.

- **: تكنولوجيا المنشأ والشكل المعماري.**

يمثل المنشأ الجزء الأساس في العمارة، والذي يمتاز بكونه مفهوم عام يجسد اتجاه القوى ومساراتها. فالمنشأ كيان ديناميكي يتغير زمنياً متطوراً من شكل إلى آخر، وأحياناً تحدث قفزات تطورية يظهر فيها أنموذجا جديداً مختلفاً عن الأول، والمنشأ الجديد يخضع لمفهوم التطور المستمر أي أن هنالك أجيالاً تكنولوجية، وبعبارة أخرى يمكن اعتبار المنشأ – وفقاً للتطورات التكنولوجية – كيان بشري الصنع يمثل مستوى إبداعي عالي في مراحل زمنية معينة وان هذا

الكيان يتطور باتجاه أعلى خدمة للمستخدم بأعلى كفاءة ممكنة وبما يلبي الحاجة المستمرة (ذات المتغيرات المتلاحقة زمانياً ومكانياً).

- **الوظائف الأساسية للمنشأ:-**

أ : **الوظيفة العملية الأساسية للمنشأ:** والتي تتمثل بنقل القوى وإيجاد الموازنة المنشئية للمبنى بحيث يكون تركيز الشكل المنشئي على عكس الانتقال الطبيعي للقوى بأقل مسارات ممكنة.

ب : **الوظيفة التعبيرية للمنشأ:** بالرغم من الوظيفة العملية البحتة للمنشأ إلا أن هذا لا يمنع من دوره في تعميق المعاني الجمالية، حيث تبرز أهمية المعاني الجمالية التي يمكن أن يضيفها المنشأ للتشكيل العام.

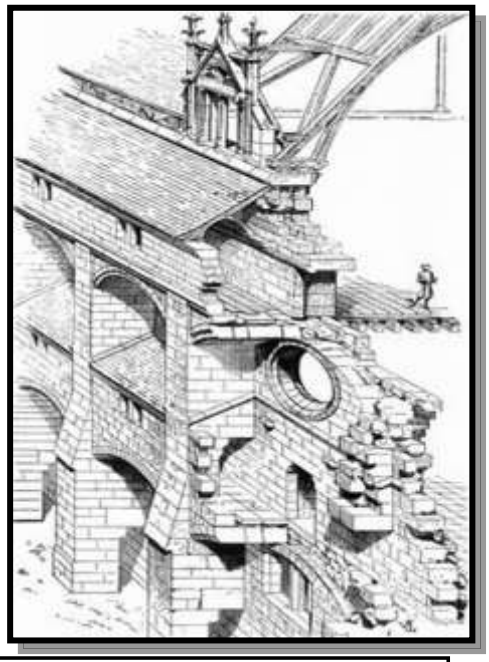
- **أنواع النظم المنشئية:**

كشفت العديد من الدراسات بان تحديد النظام المنشئي يعتمد الكثير من المحددات منها خصائص المواد وطبيعة الأساليب التنفيذية المتاحة ونوع المتطلبات الاجتماعية المراد تلبيتها فضلاً عن طبيعة الفعاليات الوظيفية وغيرها.

ومن ابرز التصنيفات للنظم المنشئية هو تصنيف (Angerer):-

أ: المنشآت الصلدة (المصمتة) Solid Structures :

تعّد من اكثر النظم انتشاراً خاصة في العمارة التقليدية المحلية لارتباطها بشكل وثيق بالأساليب التقليدية في الإنشاء. بالتالي فان مركباتها الأساسية الحاملة للأثقال تعتمد على تراكب الوحدات البنائية، وقد اتبع هذا النظام المنشئي في النماذج المعمارية التقليدية لملاءمته للمواد الطبيعية المتاحة، ويستغل المصمم أحياناً أبعاد الوحدات البنائية المكونة للنظام وتراكبها للحصول على معالجات متنوعة لما تتيحه تلك الأبعاد من إمكانيات التنويع في أساليب الربط والتركيب بحيث توفر إمكانيات منشئية وتزيينية متعددة. [٨/ص:٤].شكل(١).



شكل رقم (١) المنشأ المصمت

بحيث يمكن القول بان "الحد الخارجي للمبنى هو ذاته الجدار الخارجي"، فهنالك تكامل من خلال عناصر تعريف الفضاء أحدهما يدور باتجاه العالم الخارجي والآخر يدور باتجاه الفضاء الداخلي وما بين هذين الجانبين تتواجد كتلة تمثل الجسم الفعال لهذا النظام المنشئي. [٨/ص:٥].

- **الجدران الحاملة والسمك الفعال:** يشكل السمك الفعال العامل المهم في تحديد السمات الفيزياوية للجدار حيث يمثل هذا البعد في الجدران الحاملة الجانب الذي يفصل بين منظومتي القشرة الداخلية والخارجية مما يعطيه الأهمية الكبيرة كجانب منشئي وتشكيلي ساهم في إغناء العمارة العراقية تعبيرياً من خلال المعالجات الداخلة عليه.

وتكمن أهم المؤثرات التي من شانها التأثير في تحديد السمك الفعال للجدران الحاملة بالمؤثرات التالية:-

* **طبيعة الأحمال المنتقلة ونظام التسقيف المعتمد:** والذي ساهم في تحديد

ابرز السمات الشكلية والمنشئية للجدران الحاملة. حيث كان في التغير في انتقال الأحمال من كونها متمركزة (Concentrated Loads) بنقاط محددة، إلى كونها موزّعة (Distributed Loads) وعلى طول المساحة السطحية للسمك الفعال للجدار الأثر الكبير في تحديد السمك الفعال.

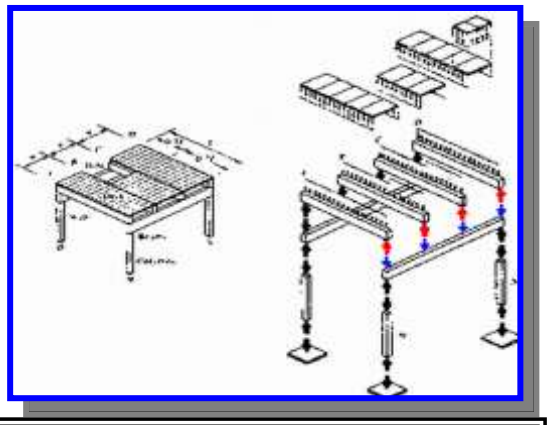
* **طبيعة المادة البنائية المعتمدة وتطوراتها النوعية:** مثّل هذا الجانب ابرز محددات السمك الفعال للجدران الحاملة وعلى مر المراحل.

ب: المنشأ الهيكلي Skeleton Structure :

يتكون هذا النظام من عناصر منشئية واضحة ومتميزة تتحمل الأثقال وتقاوم القوى المؤثرة على المبنى، بالرغم من أن تلك العناصر لا تعمل على تشكيل الغلاف البيئي المحيط بالفضاء (الفصل الواضح بين منظومتي المنشأ والقشرة الخارجية). ويعتمد هذا النظام على العناصر والمواد الخطية (Linear Materials) لنقل الاجهادات خلال شبكة من العناصر المستقلة والمفصولة فيزيائياً عن مكونات منظومة القشرة الخارجية التي تفصل ما بين الداخل والخارج الأمر الذي يحرر العناصر المنشئية من مهمة العزل البيئي. [٩/ص:٢٢]. شكل (٢).

كما يستند هذا النظام في عمله على الحمل المنقول ذي الطبيعة النقطية (Pointed) مما يساهم في توفير مرونة تصميمية اكبر في تشكيل حدود الفضاءات الداخلية والخارجية بمعزل عن المحددات المنشئية فضلاً عن تقليل المساحة المطلوبة للعناصر المنشئية. شكل رقم (٣).

شكل رقم (٢) المنشأ الهيكلي لمبنى Court Building للمصمم Martine Kenny ٢٠٠٤



شكل رقم (٣) توزيع القوى في المنشأ الهيكلي

ج: المنشأ السطحي Surface Structure :

يمتاز باعتماده على عناصر بنائية سطحية ذات سمك صغير عملت على نقل القوى والأحمال وتحديد الفضاء. والتي يتم من خلالها توزيع الأحمال والاجهادات في الفراغ (في الاتجاهات الثلاثة) وليس في اتجاه واحد كما في الأنواع السابقة، كما يمتاز بامتلاكه حالة التطابق بين المنشأ والفضاء الداخلي. شكل(٤).

شكل رقم (٤) المنشأ السطحي لمبنى Congress Hall للمعمار 1956 Hugh Stubbins

- **المـــتطلـ**

يات المنشئية Structural Requirements

أ : الوظيفية المنشئية Structural Functionality : يجب أن يتوافق المنشأ قدر الإمكان مع متطلبات الفضاء الوظيفية إضافة لمتطلبات الحركة والإسناد التي تفرضها تلك الوظيفة. وعليه فالمتطلبات الوظيفية تلعب الدور الكبير في نوع المنشأ وتقرير شكله المطلوب، بما تحويه من تأثير قياس البحر وموقع نقاط الدعم والإسناد المطلوبة ومقدار الأحمال والتكوين الضروري للفضاء الداخلي. [١١/ص:١٩٤]. وهكذا فان المتطلبات الوظيفية ممكن أن تساهم أحياناً في اقتراح الشكل المنشئي الملائم. بحيث يمكن أن يكون الشكل والمنشأ هنا متوائم مع الوظيفة شكل(٥).

شكل رقم (٥) دور الوظيفة في المنشأ في مبنى الباوهاوس للمعمار كروبيوس ١٩٢٦

ب : الاستقرار المنشئي Structural Stability:

يجب أن يأخذ الشكل المنشئي بالحسبان القيود المفروضة عليه من قبل طبيعة الموقع المختار وبكل ما يتضمنه من طوبوغرافيا وطبيعة التربة والصخور والخدمات التحتية إضافة للقيود البيئية. بحيث يصل إلى الاستقرار المنشود ومن ابرز المبادئ المرتبطة بالاستقرار المنشئي ما يعرف بـ:

* **مركز الثقل المنشئي (Structural Center of Gravity)/** وهي نقطة خاصة تتوزع حولها جميع كتل المنشأ بصورة متساوية وبالتالي الوصول إلى حالة الاستقرار.

كما أن البعض يشدد على ضرورة التكامل بين مركز الثقل المنشئي مع مركز الثقل البصري بحيث يكون المبنى في حالة قبول بالنسبة للمتلقي متى ما كان هنالك ذلك التوافق بين مركزي الثقلين. شكل (٦)

شكل رقم (٦) مركز الثقل المنشئي وعلاقته بمركز الثقل البصري لمبنى City Hall للمعمار Norman Foster 2002

ج: الموازنة المنشئية Structural Equilibrium: ويكمن المتطلب الأساسي للموازنة في الضمان بان المبنى كلاً أو أيٌ من أجزائه سوف لن يتحرك. ولكن يجب ملاحظة عدم القدرة على تنفيذ هذا المتطلب بشكل كامل لان بعض الحركات لا يمكن تجنبها بالإضافة إلى كونها ضرورية في بعض الأحيان.

د: المتانة المنشئية Structural Strength:

المنشأ هنا يجب أن تكون لديه الدرجة الكافية من المقاومة والثبات ضد الأحمال المسلطة وان لا تظهر عليه أي علامة من علامات الضعف بسبب الإجهاد المفرط أو الحركات الغير متوافقة للمواد والمركبات.

هـ: الاقتصادية المنشئية Structural Economy:

بالرغم من أن الاقتصاد لا يكون متطلب معماري خاصة في المنشآت ذات الأغراض الرمزية والتذكارية. ولكن السمة النفعية للمنشأ تعد من الأمور الأساسية حتى في الأنظمة المنشئية للمباني ذات الصفة غير النفعية. [١١/ص:١٩٥]. وعليه فالمنشأ الذي يتم انتخابه يجب أن يمتلك إمكانية الإنشاء والتركيب من خلال استخدام تقنيات إنشائية موجودة ومتوفرة وبشكل كفوء إضافة إلى توفر مكونات المنشأ المادية بشكل اقتصادي وبالكميات المطلوبة.

و: الجمالية المنشئية Structural Aesthetics:

تأثير الناحية التعبيرية على المنشأ لا يمكن أن يُنكر أو يُهمل بسبب فرض الكثير من الأفكار والعقائد الجمالية على المنشأ. حيث يكون المنشأ هنا من أهم عناصر الوصول للتصميم الناجح، كما أن حجم ذلك التأثير يختلف باختلاف النظرة لأهمية التأثير المنشئي وباختلاف حجم ومقياس المبنى.

☒ الكفاءة والإبداع المنشئي Structural Efficiency & Structural Creativity:

يمكن تعريف كفاءة المنشأ بأنها النسبة بين الحمل الكلي إلى الوزن الميت فكلما قل الوزن الميت ازدادت الكفاءة المنشئية. **وبالتالي يمكن استخلاص ابرز إمكانيات تحقيق الكفاءة المنشئية وفق ابرز الطروحات النظرية:**

١. من خلال استيفاء المتطلبات المنشئية الأساسية السابق ذكرها للحصول بالنتيجة على المنشأ الكفوء (والذي سماه Salvadori بالمنشأ الأفضل). [١٣/ص:٧٦].

شكل رقم (٧) مشابهة المنشأ للطبيعة لمبنى قصر العمال للمعمار Nervi

٢. الأشكال المنشئية المبدعة هي التي تكون مشابهة لتلك الموجودة في الطبيعة (سواءً على مستوى الشكل أو المضمون). لان تلك الأشكال يجب أن تكون بالضرورة كفوءة لأنها بقت واستمرت. [١١/ص:٢٠٨]. شكل (7).

٣. أما النظرة الحديثة لمقياس الكفاءة المنشئية فتشدد على أن الشد Tension كمنظومة منشئية تعد من أهم عوامل تحقيق تلك الكفاءة لأنها اكثر كفاءة لنقل الأحمال. فالشد هو أحد الوسائل الأكثر كفاءة من ناحية انتقال الأحمال وتحمل الاجهادات بالرغم من أهمية عناصر الإنضغاط فيه. [١١/ص:٢٠٩]. شكل(٨).

شكل رقم (٨) منشآت الشد / Millennium Bridge / للمعمار Foster 2001

٤. يمكن أن يكون الإبداع المنشئي في جانبين: الأول هو الانسلاخ والتحرر من الألفة والترابط مع تجارب الإنجازات السابقة، والثاني في الفهم المتعمق والموضوعي للفعالية المنشئية والذي يعطي بدوره العمق الضروري لمعنى الخبرة ويجعل منها ممكنة التحرك خلفها.

٥. ينطلق المنتج المبدع من (الأسلوب الصادق) والذي تكمن خصائصه الضرورية في (الجوهر المنشئي– الغياب الضروري للتزويق والزخرف– نقاوة الخط والشكل). [١٢/ص:٢٦].

☒ عناصر المفردات المنشئية في التشكيل المعماري:

يمكن إدراج المفردات المنشئية في التشكيل والتكوين البصري في جانبين:

أ- **مادي :** خاص بمواد البناء وأنظمة البناء والتشكيل: و هو يقترن بمجالي العلم والصناعة فيأخذ عنها المواد والتقنيات الحديثة وهو في هذا الشق يخاطب المستوى العقلي للإنسان.

ب- **رمزي :** خاص بالمعتقدات ، فيخاطب المستوى الروحي للإنسان من خلال تعبيره عن المعتقدات والشعائر الخاصة بالمجتمع خلال العملية البنائية. [٧/ص:html].

- **الشكل المعماري Architectural Form:**

الشكل (form) هو ترتيب معين للعناصر أو المفردات المفصولة عن بعضها بمسافات (intervals) زمانية أو مكانية محددة متخذة هيئة معينة. وهو يتألف من عناصر فيزيائية (هي الكتلة) تحيط بعناصر غير فيزيائية (هي الفضاء). وعليه فالشكل هو تمثيل مادي بصري ناتج عن تفاعل جملة من المتطلبات الاجتماعية والتكنولوجية لتأسيس مادة قابلة للإدراك تمتاز عن غيرها بكونها أماكن للإيواء الإنساني، فالشكل المعماري لم يكن أبداً البنية الظاهرة وإنما هو البنية المدركة. وتعتمد إقامة الشكل على توفر الإمكانيات المنشئية القادرة على تحويل مجموع الأفكار إلى واقع مادي محسوس ويكون الشكل الفيزيائي بعد تنفيذه هو وسيلة الاتصال ونقل المعاني بين المعمار والمستعمل.

- **إمكانيات تكامل العلاقة بين الشكل المنشئي والشكل المعماري:**

المقصود بالتكامل هو العلاقة التي ترتبط بها المنظومات الرئيسية المكونة للمبنى، بشكل لا يجعل منها منظومات مستقلة الواحدة عن الأخرى، **فالتكامل كمفهوم تكنولوجي** يكمن بالتوازن والانسجام بين وظائف المبنى مع مكونات المنشأ والأنظمة الأخرى. [٦/بحث منشور]. وتبرز هنالك عدة إمكانيات للوصول لذلك التكامل منها:

أ) من خلال ترجمة للقرارات المتجانسة والمنسقة للجهات ذات العلاقة بالمبنى. وهذا الانسجام بين المنشأ والشكل المعماري تعد المفتاح للنجاح التعبيري للمبنى. [٥/بحث منشور].

ب) من خلال التاكيد على كون العمارة فن تنظيم الفضاءات، ومعانيها التعبيرية تكمن في المنشأ. فالإحساس بالتكامل وفق هذا المنظور يأتي من الكمال التقني. [١/ص:١٣].

ج) بتحقيق (الصراحة المنشئية)، حيث أن "كل تحسين في الوظيفة والفعالية التكنولوجية أو التقنية للنتاج يقابله تحسيناً في نوعيته التعبيرية والجمالية". [١٢/ص:٢٦].

د) بإعطاء أهمية للأسلوب الذاتي (ذاتية المصمم) في التعامل التصميمي للوصول إلى التكامل من خلال إمكانية التلاعب بأشكال العناصر المنشئية ذات الطبيعة العملية الصرفة وذلك من خلال تحويل المتطلبات المنشئية إلى عناصر جمالية مع عدم الممانعة من إخفاء بعض الحقائق المنشئية لتحقيق التعبير والجمالية المطلوبة.

- **تطورات العلاقة التأثيرية بين المنشأ والشكل في العمارة العراقية خلال القرن العشرين:**

مرت عمارة القرن العشرين في العراق بمتغيرات كثيرة امتازت بسرعتها وتفاوتت في شدة تأثيرها على مر القرن بحيث حمل كل عقد حدثاً مؤثراً من شأنه ان يحدث أثراً مهماً في نتاج العمارة وعلى مختلف المستويات فقد امتازت

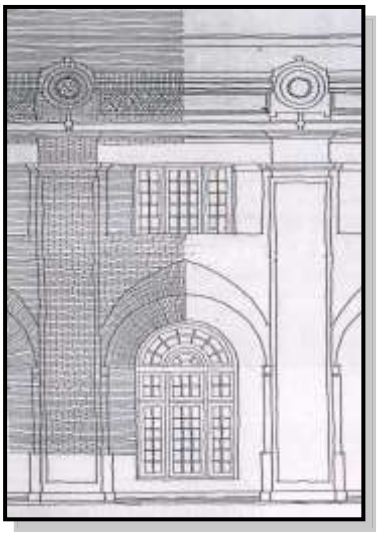
عمارة القرن العشرين في العراق بظهور عدة طرز أو أساليب في الممارسة المحلية حيث تم تصنيفها استناداً الى توجهاتها التكنولوجية خاصة ما يتعلق منها بالجانب المنشئي،

حيث برزت هنالك **أربعة مراحل رئيسية:**

أ : المرحلة الأولى: (١٩٠٠-١٩٤٠) "شيوع محددات المنشأ المصمت":

تمثل هذه المرحلة تجسيداً للعمارة التقليدية المحلية في استغلال التقاليد والموروث على مستوى المنشأ والتشكيل والإنشاء وبما ينسجم مع المعطيات التي يفرضها الواقع العملي المحلي من نواحي العمالة والتقنيات (في الإنتاج والتنفيذ) والمواد التي سادت في تلك الفترة التي امتازت بتنوع مؤثراتها لما مرت به من تعدد في مصادر تأثيرها وخصوصية كل مصدر.

وبالنتيجة فالنظم المنشئية المعتمدة في تلك المرحلة اعتمدت إجمالاً على النظم والمركبات المنشئية التقليدية المعروفة والشائعة في المنطقة كالجدران الحاملة السميكة والتسقيف بالأقبية أو العقادة كما اقتصرت على المادة البنائية الشائعة (الطابوق) في تنفيذ تلك النظم والمركبات المنشئية. كما امتازت منشآت تلك المرحلة باعتمادها الكلي على مواد النظام المنشئي نفسها لإحراز التأثيرات الفنية والتشكيلية على واجهات المباني، فالطابوق المستخدم في المنشأ هو ذاته الذي يعطي قوى التعبير المعماري المتأتي من ملمسه ولونه وصفاته، شكل رقم (٩). حيث مثّل التطور النوعي لتلك المادة تحولاً نوعياً في جانِبيّ المنشأ خاصةً والعمارة عامةً لما ساهم به من رفع الكفاءة المنشئية للجدران الحاملة وبالتالي التأثير على السمك الفعال للجدران ومن ثم التقليل من السمك المطلوب وما ترتب عليه من زيادة المرونة في الفتحات والتقليل من المادة المستعملة.



شكل رقم (٩) جدار الشعبة الدينية / ويلسون ١٩٢٢

مما جعل المنظومة المنشئية هنا أحد الأعمدة الرئيسية في التكوين العام. بحيث جاءت اغلب التكوينات نابعة من الشكل المنشئي للمركبات المنشئية.

كما نلاحظ ذلك في مبنى الشعبة الدينية (جامعة آل البيت) حيث كان هنالك استغلال لما وفرته تلك المركبات من كفاءة منشئية وتعبيرية. وبالتالي تكاملية العلاقة بين المنظومة المنشئية ومنظومة القشرة الخارجية في خلق التكوينات والتشكيلات المطلوبة (النابعة من طبيعة النظام المصمت والمادة البنائية). بحيث جاءت اغلب تكوينات المبنى نابعة من الشكل المنشئي للمركبات المنشئية (الشكل يتبع انتقال ومسارات القوى). شكل رقم (١٠).

شكل رقم (١٠) الرواق الشعبة الدينية / ويلسون

ب : المرحلة الثانية: (١٩٤٠- ١٩٦٥) "شيوع المنشأ الهيكلي":

تجسد مرحلة العمارة المتعاطفة مع النزعة الحداثوية والمقلدة للعمارة الغربية بكل تقنياتها وأشكالها ذات الصبغة العالمية (لمجاراة الكفاءة والادائية الوظيفية والمنشئية العالمية). حيث كان لظهور الأساليب والنظم المنشئية الجديدة وخاصة النظم الهيكلية باستخدام الحديد والخرسانة المسلحة والركائز العميقة الدور الكبير في خلق تحولات نوعية معمارية كبيرة (خاصة على مستوى الأبنية العامة والإدارية ذات المقياس الكبير نسبياً).

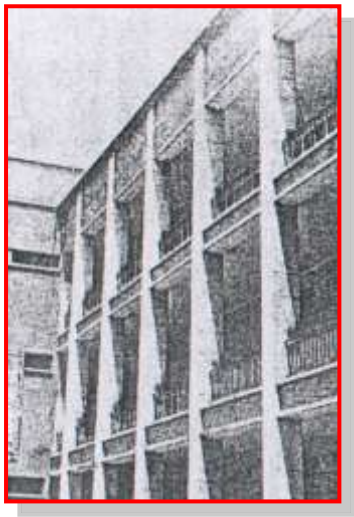
حيث أشرت تلك المرحلة حالة فصل المنظومة المنشئية عن منظومة الشكل الخارجي لتعطي بذلك مرونة تصميمية اكبر الأمر الذي أتاح حرية اكبر في التصميم والتخطيط. كما أن تلك التطورات أعطت بالمقابل إمكانيات كبيرة في زيادة عدد الطوابق.

وعليه يكمن دور النظام المنشئي هنا في تحقيق الشكل النهائي للفضاء دون افتعال أو إضافة عناصر غير مطلوبة. وظهور ما عُرف بـ(الزخرف الجديد) النابع من الشبكة المنشئية (Structural Grid) بعناصرها العمودية والأفقية والمؤثرة مباشرة في الشكل المعماري سواء أكانت تلك الشبكة ظاهرة أم مخفية. شكل(١١)، حيث ساعد النظام المنشئي الهيكلي هنا في الكثير من جوانب الفعالية التصميمية والبنائية وخاصة على مستوى الفضاء من خلال إعطاء مساحات واسعة بين العناصر المنشئية الساندة الأمر الذي أدى بالمقابل إلى زيادة واسعة في الفضاءات الداخلية وزيادة المرونة التصميمية فيها وصولاً إلى تحقيق المخطط المفتوح. وهذا نابع أساساً من الإمكانيات المنشئية للنظم الهيكلية من ناحية انتقال القوى فيها حيث يعتمد هذا النظام على المركبات الخطية لنقل الاجهادات خلال شبكة من العناصر المستقلة والمفصولة فيزيائياً عن مكونات منظومة القشرة الخارجية التي تفصل ما بين الداخل والخارج الأمر الذي يحرر العناصر المنشئية من مهمة العزل البيئي، وتتطلب تلك العناصر استخدام مواد كفوءة في تحملها لاجهادات الشد والانضغاط معاً.

شكل رقم (١١) مبنى الساعاتي / قحطان عوني

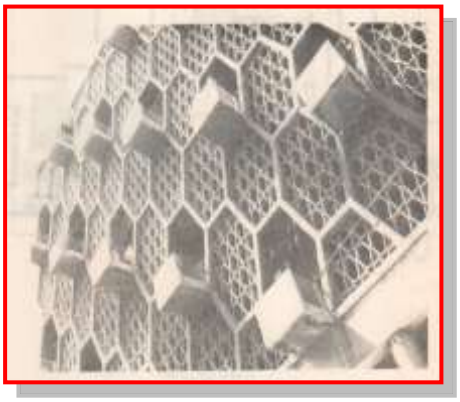
شكل(١٢).

فالتشكيل بالنتيجة كان يعبر عن نمط النظام المنشئي ولكنه لا يظهره بالضرورة، فالشكل هنا يميل إلى عكس الوظائف، والوظيفة هنا يمكن أن ترتبط بالخصائص الوظيفية للعناصر المنشئية وطبيعة المواد البنائية وحتى طريقة الإنشاء. بالرغم من شيوع المعالجات الشكلية ذات المنحى التقليدي كما في مبنى المدرسة الجعفرية للمعمار جعفر علاوي، شكل(١٢). حيث تم استخدام الجدران الطابوقية المكشوفة ذات الوظيفة غير المنشئية (الجدران غير الحاملة للأثقال) والتي استندت كلياً على المنشأ الخرساني الهيكلي الذي اخذ على عاتقه إسناد الأرضيات الخرسانية. وهذا مثّل توظيفاً للـ(النظم الهجينة) التي جمعت بين تعبيرية الشبكة المنشئية للنظام الهيكلي وبين تعبيرية الجدران الطابوقية بالرغم من كون الأخيرة جاءت بوظيفة غير



شكل رقم (١٢) المدرسة الجعفرية/ جعفر علاوي

منشئية. وفي جانب آخر ظهر لدينا بعض المعالجات التي تعكس انعدام الصراحة المنشئية من خلال استخدام مادة الخرسانة في الواجهة بتشكيلات تعكس القوة والصرامة على الرغم من كونها تقوم بوظائف غير منشئية أصلاً كما هو الحال في واجهة مبنى خان الباشا الصغير للمعمار عبد الله إحسان كامل. شكل(١٣).



شكل رقم (١٣) مبنى خان الباشا الصغير / عبد الله إحسان كامل ١٩٥٦

ج: المرحلة الثالثة: (١٩٦٥- ١٩٩٠) "الجمع بين التراث والتقنيات المنشئية الحديثة" :

امتازت بكونها مرحلة السعي نحو التعاطف مع التراث بأسلوب يعكس التقدم التكنولوجي بحيث تكون المحاولة لاستلهام التراث كوجود ثقافي أو فيزياوي وتحويره أو تجريده وبما يلائم التوجهات التكنولوجية والتكوينية المعاصرة. فالتركيز كان هنا على التشكيل العام النهائي منطلقاً وفق مبدأ توفير الإسناد والدعم له من قبل المنشأ دون الاعتماد على نمط راسخ. حيث كان الهدف يكمن في المحاولة لإعادة العلاقة بين الشكل والمعاني بتوظيف عدد من المبادئ الجمالية والأنظمة التقنية بما فيها الأنظمة المنشئية الحديثة عن طريق اعتماد ثنائية تجمع ما بين القديم والحديث بغية إغناء المعاني. فطبيعة العلاقة بين كل من المنشأ والشكل في عمارة هذه المرحلة هي علاقة ضمنية، والشكل هنا متنوع الطرز والأنماط يستخدم المعالجات الزخرفية التي تخفي الجوهر المتمثل بأنماط النظم المنشئية التي قد تأتي كتحصيل حاصل للمعالجات الشكلية. [٣/ص:١٩].

شكل رقم (١٤) اتحاد الصناعات للمعمار رفعة الجادرجي

وبالتالي استخدام مفردات التراث المعماري في التشكيلات التكوينية الفنية للمباني المنفذة وتوظيفها في المنشأ الحديث. بحيث ظهرت لدينا محاولات تهدف إلى بلورة العمارة العراقية المحلية وإيجاد لغة جديدة تجمع العناصر والعلاقات التراثية مع التقنيات المنشئية الحديثة ومتطلباتها (خاصة توجهات ما بعد الحداثة) كما في أعمال محمد مكية ورفعة الجادرجي وقحطان عوني وغيرهم رغم المبالغة في هذا المنحى بعض الأحيان.شكل(١٤).

بالتالي فقد مثّل التشكيل المعماري حالة من الانفصام الضمني بين الشكل الظاهر مع مضمون المنشأ، فالشكل الخارجي للمبنى ليس مهمته عكس مضمون فعاليات المبنى ووظيفته بقدر ما هو إقرار مسبق بحيث كانت بعض المشاريع تمتاز بالتضاد في معالجات الواجهات التي كانت معقدة بشكل غير عادي وبين وضوح وبساطة المخططات والهيكل المنشئي الواقع خلفها. [٢/ص:html]. فالنظام المنشئي ساهم هنا في خلق هيكل تستند أو تعلق عليه الستارة التي تشكل التكوين العام لواجهة المبنى، شكل(١٤).. كما في اغلب أعمال المعمار رفعت الجادرجي.

د: المرحلة الرابعة: (١٩٩٠ - ٢٠٠٠) "العودة للتقاليد المنشئية/بروز اللاتكتونية المنشئية:

تمثل مرحلة الخلط واستنساخ الأشكال والأنماط من عدة توجهات (التراثية والمعاصرة) على المستوى الشكلي والتكويني وحتى على مستوى التفاصيل والمواد باستخدام تقنيات وأساليب تقليدية محلية في المنشأ.

مما افرز ازدواجية في التعامل والاهتمام على مستوى الأشكال والمنشأ والمواد تبعاً لاختلاف المتطلبات التي يمليها رب العمل أو الواقع التنفيذي. حيث مثل المنشأ هنا جانب المنظومة التي توفر المتانة والإسناد مستغلة المواد المنشئية التي تتيح تحقيق مختلف التكوينات والتشكيلات بالرغم من اتباع التقنيات التقليدية في إنشائها ومن ثم استخدام مختلف مواد التغليف في خلق القشرة المغلفة لخلق الشكل النهائي الناتج. ولو على مستوى المقياس الصغير حيث كان السائد هنا عمارة المنازل والعمارات التجارية.

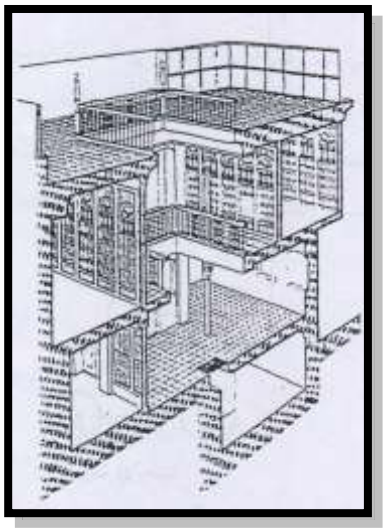
أما الأهداف فقد كانت لا تعدو عن كونها تفخيمية رمزية بعيدة عن الأسباب المنشئية أو الوظيفية، حيث أصبحت من الضرورات التصميمية المفروضة من قبل رب العمل. وهذا النهج يعد مؤشراً واضحاً على انتهاج اللاتكتونية في التعامل التصميمي.

شكل رقم (١٥) عمارة المرحلة ٤

وبالنسبة لدور المنشأ في المباني ذات المقياس الكبير نسبياً كالعمارات التجارية والفنادق المتوسطة وغيرها فقد كان النظام الهيكلي يمثل النظام المنشئي الأساس إن لم يكن الوحيد لإنشاء مثل هكذا مبانٍ، والتي جسدت حالة من التكرار في الأبعاد والمواد والمعدات المستخدمة في الإنشاء بغض النظر عن المحددات والحسابات المنشئية التي تفرضها طبيعة المشروع الوظيفية وأبعاد فضاءاته أو مدى القوى والأحمال التي يتعرض لها، كما أن تلك النمطية في استخدام الشبكة المنشئية قد أثرت وبشكل لا يقبل الشك في تحديد الأشكال والتصميمات الناتجة حيث مثلت المحددات المنشئية هنا المؤثر الأكبر والمقيد نحو المرونة التصميمية، وان تلك المحددات المنشئية كانت نابعة من الثقافة المنشئية المحدودة التي أدت بالنتيجة إلى تلك النمطية أو التكرار المفرط. وبالتالي ظهرت التكوينات عبارة عن مواد مضافة أو ملصقة على الهياكل المنشئية لأغراض جمالية تزيينية بحتة وليس بأسلوب استخدام عنصراً يحقق وظائفها المصممة لها. وهذا ما يؤكد من تركيز نتاج المرحلة على الجانب التكويني الاظهاري دون الاهتمام بالجانب المنشئي. شكل(١٥).

☒ أنماط طبيعة دور المنشأ في الشكل وتطبيقاته على العمارة العراقية:-

أ: المنشأ هو الشكل (توحد العلاقة وتكاملها):



شكل رقم (١٦) البيت التقليدي

يتركز هذا الجانب في استخدام نمط منشئي بارز تتوحد فيه كل من المنظومتين المنشئية والقشرة الخارجية، وهذا ما برز جلياً في تكوينات عمارة **المرحلة الأولى** حيث جاءت التكوينات بشكل عام تماثلية لمعظم تصميمات المباني المنفذة آنذاك بسبب محددات ثبات المنشأ كنظام وتأثيرات المادة البنائية مما فرض نمطاً محدداً من التكوينات على المباني بغض النظر عن الجانب الوظيفي. حيث تميز النتاج بكونه عمارة ذات نزعة تكاملية من ناحية العلاقة والتأثير بين المنشأ كمنظومة إسناد أساسية وبين القشرة الخارجية كمنظومة فضائية تكوينية تعطي للمبنى شكله المميز له بحيث يصعب بمكان فصل المنظومتين عن بعضهما. فضلاً عن صعوبة التفريق بين المنشأ كمفهوم إسنادي وبين الإنشاء كتجسيد مادي لهذا المفهوم. وهذا ينصب بدوره في توحد العلاقة بين المنشأ والشكل من نواحي الاستجابة والتأثير لامتلاكهما خصائص مشتركة في جوهرهما، بحيث لا يمكن التفريق بينهما فأحدهما يعتمد على الأخر ويكوّن الآخر. شكل(١٦). وبالتالي فهنالك تأثير متبادل لكل منهما في الآخر، فعندما يكون هنالك اهتمام بالجانب التعبيري للشكل المعماري يدفع هذا باتجاه صياغة ومعالجة الشكل المنشئي على مستوى المواد والتفاصيل للوصول للتعبيرية المطلوبة لان المنظومتين مشتركتان بنفس الصفات الفيزياوية والتعبيرية البصرية. وبالتالي فالجدار المصمت جسد هنا المنشأ والقشرتين الداخلية والخارجية معاً.

فالمحددات المنشئية على مستوى متانة المنشأ ككل أو المواد التي تشكله تمثل إحدى عوامل التشكيل أن لم تكن أقواها. ولهذا نجد الكثير من المركبات المنشئية استندت أشكالها على طريقة توزيع ومسار القوى المنتقلة خلالها كما نجد ذلك واضحاً في العقود المتقاطعة والأقواس وغيرها.

ب : المنشأ وسيلة لإنتاج الشكل المعماري (جزء مكمل للتشكيل):

لا يشكل المنشأ هنا عمارة بمفرده وإنما يجعلها ممكنة الحدوث حيث يوفر البنية المادية ذات الطابع الإنساني، فيجسد المنشأ هنا الوظيفة العملية (المادية) له. ويمكن ملاحظة ذلك التوجه في تطبيقات عمارة **المرحلة الثانية**، وما رافقها من تكرار لغة اغلب المعالجات الشكلية على الرغم من التطورات المنشئية الكبيرة الداخلة على الفعالية التصميمية وهذا خلق نوع من عدم الصراحة المنشئية في التعبير الناتج.

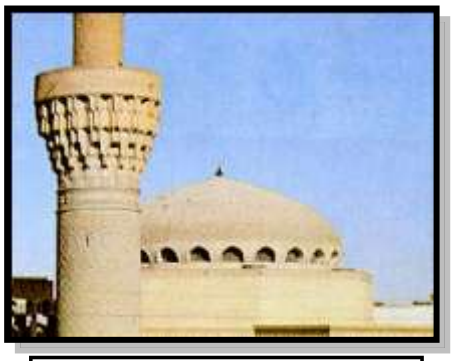
شكل رقم (١٧) مبنى كمال السامرائي / قحطان المدفعي

فالأشكال هنا كانت مجردة تحكمها الوظائف وتحددها الأنماط المنشئية التي اعتمدت على الشبكة المنشئية المتعامدة والمخطط المفتوح الذي انعكس بالمقابل على الأشكال التي جاءت على درجة من التجريد والبساطة. شكل(١٧). والتركيز كان على السياق الوظيفي والشبكة المنشئية كمصادر لخلق الأشكال التي تدل على معاني محددة بمتطلبات تلك الوظيفة أو المنشأ. [١٠/ص:٨٧].

ج : المنشأ أحد مبادئ إنتاج العمارة:

يكون المنشأ هنا مرتبطاً بالعمارة والتشكيل النهائي من خلال توفير الإسناد الملائم للمبنى وما يساهم به كمنظومة إسنادية. وهذا الجانب من التأثير المنشئي يمكن تطبيقه على نتاج **المرحلة الثالثة** والتي تميزت –ولو بشكل جزئي– بأنها ذات خصائص ثنائية من خلال استعمالها للعناصر التقليدية (باعتبارها أدوات اتصالية) بوسائل تنفيذ غير تقليدية. فالعمارة هنا ركزت جل اهتمامها على الشكل، أما اختيار نمط المنشأ فيها فيكون منطلقاً من ذلك الشكل وفق مبدأ توفير الإسناد والدعم له دون الاعتماد على نمط موحد. فالشكل هنا متنوع الطرز والأنماط يستخدم المعالجات الزخرفية التي تخفي الجوهر المتمثل بأنماط النظم المنشئية. [٣/ص:١٩]. فظهرت ضمن منحى الاستلهام التراثي ثلاثة أنماط أساسية على المستوى الشكلي وما رافقه به من تأثيرات منشئية وهي:

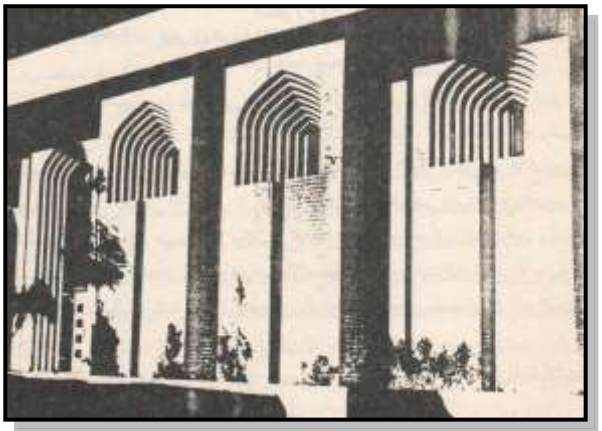
أ– التوظيف المباشر للعناصر الشكلية التراثية والذي يمثل مرحلة التعاطف مع التراث بحيث يكون هنالك تداخلاً بين العناصر والتشكيلات التراثية مع المركبات المنشئية الحديثة لتكون الحصيلة عمارة تستغل مجالات التقنية والإسناد المنشئي ولكنها متكيفة كلياً مع الجانب المحلي. [٤/ص:٢٥٥]. شكل(١٨).



شكل رقم (١٨) جامع الخلفاء للمعمار محمد مكية

ب– توظيف تشكيلات التراث المحلي بالأسلوب الذي يوازن بين المباشرة والتجريد، فالطابع العام للمنشأ يبقى عمارة دولية بكل تقنياته وخواصه عدا تلك الإدخالات الشكلية المستنسخة، فالتعبير هنا يكون بطرز معاصرة في المنشأ وبعض التكوينات لتشكيلات تقليدية من التراث. شكل(١٩).

ت– توظيف العناصر التراثية بعد تأويلها وتجريدها إلى حد بعيد (التراثية التجريدية) وهذا النهج يعتمد بدرجة كبيرة على المادة التقليدية المحلية، كالطابوق المكشوف والعناصر المنشئية المكشوفة أيضاً مع استخدام مادة الخرسانة في إنشاء بعض العناصر التراثية. شكل(٢٠).



شكل رقم (١٩) مصرف الرافدين/الكوفة للمعمار محمد مكية

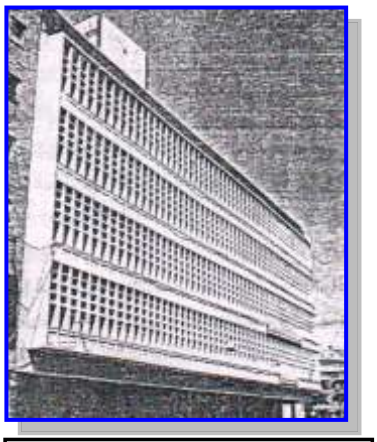
شكل رقم (٢٠) شركة التأمين / رفعت الجادرجي ١٩٦٦

كما يمكن تطبيق هذا المنحى على توجهات عمارة **المرحلة الرابعة** وما مثلته تلك العمارة من استعارة واستنساخ الأنماط والأشكال من التوجهات والعمائر المتنوعة وتعليقها وإلصاقها على المنشأ التقليدي. حيث مثّل المنشأ هنا جانب

المنظومة الاسنادية بكل متطلباتها وبتكرار كبير في المعالجات على مستوى الشكل المنشئي أو على مستوى المواد بغض النظر عن طبيعة ومقدار الأحمال والقوى المؤثرة في اغلب الأحيان، بالضد من منظومة القشرة الخارجية التي امتازت بأنها ذات صبغة التعقيد والتنوع الكبير.

د: المنشأ وسيلة لإدراك مقياس العمارة:

يعد المنشأ هو المعبر عن المقياس المدرك من قبل المتلقي، فإدراك المنشأ هو الأداة أو المفتاح لإدراك العمارة من نواحي المقياس والمعاني والرموز. وهذا التوجه يبرز في بعض نتاجات عمارة **المرحلة الثانية** وما مثلته من انتشار واسع وكبير للمنشأ الهيكلي بكل خواصه الشكلية وشبكاته المنشئية ذات الأعمدة والجسور الخرسانية خصوصاً في عمارة المقياس الكبير غير المألوفة مسبقاً. فقد عمل المنشأ في بعض الأبنية سواء بشكل مقصود أم لا على تحقيق ذلك الشعور بالمقياس الحقيقي بالنسبة للمتلقي، وقد اتخذ هذا المنحى وسائل وطرق عديدة اتبعت لهذا الغرض منها ظاهرة البالكونات أو التطليعات المستغلة لإمكانيات المادة المنشئية (الخرسانة المسلحة) ومميزاتها الفيزياوية وما أتاحته من مرونة تصميمية ومنشئية وبالتالي ظهرت الكثير من المباني حاوية على تلك العناصر والمعالجات وبأشكال متنوعة ساهمت بما لا يقبل الشك في تحقيق ذلك الإدراك الحقيقي للمقياس. شكل(٢١).



شكل رقم (٢١) مصرف الرافدين للمعمار هيرست

هـ: المنشأ أداة اللغة التعبيرية للعمارة:

برز هذا التوجه في نتاج **المرحلة الأولى** التي جسدت تطبيقاً لتلك العلاقة حيث ساهم المنشأ وخواصه الفيزياوية في خلق وتحفيز التعبيرية التصميمية، فقد مثل الجدار المصمت الخارجي ذو السمك الكبير العنصر الأكبر استغلالاً في هذا الجانب الذي أضفى مساحة جيدة للإمكانية التعبيرية المعمارية. فضلاً عن استغلال بعض العناصر والمركبات المنشئية في خلق نوع من التعبيرية. و كانت تلك التعبيرية في بعض الأحيان نابعة من الجوانب والمتطلبات المنشئية ذاتها حيث نلحظ إن هنالك معالجات شكلية لبعض المركبات المنشئية نابعة من تأثرها بانتقال ومسارات القوى خلالها (الشكل يتبع انتقال القوة).

☒ ابرز المستويات المستخلصة والمرتبطة بطبيعة الدور المنشئي في العمارة المحلية:-

أ- : المستوى الوظيفي "المنفعة والاستعمال":-

لعبت الوظيفة دوراً كبيراً في التطورات التقنية على مستوى المنشأ وخلق الشكل المعماري بالرغم من النظرة المتفاوتة لتلك الوظيفة وطريقة صياغتها وعلاقتها بالمنظومة المنشئية في المراحل المختلفة لمسار العمارة العراقية، وبتأثير مباشر من التطورات المنشئية حيث انتقلت من مستوى التقيد والمحددات في المرحلة الأولى بحيث يكون الانطلاق الشكلي هنا هو تجسيد لمحددات ومتطلبات المنظومة الوظيفية، إلى مستوى المرونة الوظيفية واعتبارها الأساس الذي ينطلق منه الشكل المعماري وتطويع المنشأ بكل إمكانياته الحديثة لخدمة هذا الغرض، وهذا ما لاحظناه في نتاج المرحلة الثانية، ومروراً بالمرحلة الثالثة التي شهدت الانفصال المتفاوت بين الوظيفة والشكل الخارجي حيث

تمثلت الوظيفة هنا بالبساطة ومرتبطة مباشرة بشبكة الهيكل المنشئي ويقابلها تعقيد شديد في المواد والمعالجات في منظومة القشرة الخارجية المكونة للمظهر العام للمبنى وبإسناد من قبل المنشأ كذلك الذي مثل الجانب الموضوعي لهما.

ب -: مستوى الخصائص الشكلية "السعي نحو مثالية التشكيل":- كان هنالك بحث دائم عن الانسجام في العلاقة المنظوماتية. وهذا الانسجام بحد ذاته كان ذا أوجه عديدة طبقاً لمتغيرات ومتطلبات كل مرحلة من مراحل مسار عمارة القرن العشرين في العراق. حيث انتقل من المعالجات الكلاسيكية والكلاسيكية المتجددة ذات التقنيات المحلية التقليدية إلى الأشكال المجردة والنقية وانتهاج منهج التخلي عن جميع الطرز التقليدية والسعي إلى إهمال الزخرفة والتزويق والتركيز بالمقابل على التجريد والبساطة (الزخرف الجديد). ومن ثم اصبح التوجه في المرحلة الثالثة لإعادة العلاقة بين الشكل والمعاني بتوظيف عدد من المبادئ الرمزية والأنظمة المنشئية الحديثة والاعتماد على ثنائية تجمع ما بين القديم والحديث. أما المرحلة (الرابعة) فكان السعي هنا في المزج القسري لمواد وأشكال وعناصر وتفاصيل من مصادر وبيئات وأماكن متنوعة لإنتاج طرز متنوعة حتى في المبنى الواحد.

ج -: مستوى تحقيق الجماليات التصميمية:- بوجود الخصائص الجديدة للمواد البنائية والأفكار المنشئية الحديثة فان إدراك الجمال شهد تحولاً وتطوراً ورافق ذلك زيادة الطلب على تلك التقنيات والأنظمة المنشئية. وبالتالي يمكن أن تطوع جميع المفردات الجمالية للعمارة العالمية وتدخل ضمن الإدراك والتقبل الجديد للفكر المحلي العراقي وهذا ما نجده جلياً في المسار التطوري للعمارة العراقية على اختلاف مراحلها بالرغم من تفاوت تأثيره.

☒ ابرز استنتاجات البحث:

١) افرز الأثر المنشئي في التشكيل المعماري - وفقاً لنمط طبيعة العلاقة -عدة مستويات خضعت لمجموعة من المحددات والمؤثرات (سواءً ذاتية أم موضوعية) على اختلاف مراحل تطور العمارة، والتي ساهمت وبشكل كبير في صياغة نمط العمارة العراقية.

٢) امتاز نتاج المرحلة الأولى بكون المنشأ فيها هو الشكل، أما المرحلة الثانية فقد كان فيها المنشأ وسيلة لإنتاج الشكل المعماري، بينما المرحلتين الثالثة والرابعة فكان المنشأ فيهما أحد مبادئ إنتاج العمارة. أما دور المنشأ كأداة اللغة التعبيرية للعمارة فقد ارتبط بنتاج اكثر من مرحلة واحدة.

٣) برز الأثر المباشر للمنظومة المنشئية في الشكل المعماري وفق علاقة (مؤثر-نتيجة) باعتبار المنشأ الجزء البالغ الأهمية في التشكيل المعماري حيث أن معظم القرارات المؤثرة في النتاج النهائي بما تتضمنه من ذاتية تصميمية مرتبط بصورة مباشرة باختيار نمط النظام المنشئي المسؤول عن تحقيق الفضاء والوظيفة والمتانة فضلاً عن التعبيرية التصميمية المبتغاة للمبنى.

٤) برزت هنالك عدة مؤثرات ساهمت في تحديد صيغ وتوجهات كل عمارة في زمان ومكان خاصة ما تعلق منها بالتكنولوجيا المنشئية، وعليه فقد برز للتكنولوجيا المنشئية جانبان أساسيان، الأول جانب فكري أيديولوجي يمثل

التطورات في النظريات المنشئية وبضمنها الحدس المنشئي والآخر جانب عملي تطبيقي مرتبط بالتطورات النوعية للمواد الإنشائية فضلاً عن التجارب والاختبارات المنشئية.

٥) كانت درجة الاستجابة والتطبيق للمتطلبات المنشئية المختلفة –على الرغم من موضوعيتها– متأثرة بالشخصية التصميمية ودرجة الوعي والتي ارتبط بدون شك بمستوى الثقافة التكنولوجية للمجتمع.

☒ المصادر والمراجع:

١) أغا، رند حازم. (٢٠٠١)، ***اثر التكنولوجيا على علاقة الشكل بالمنشأ في لغة الفضاءات الداخلية المعاصرة***، رسالة ماجستير، قسم العمارة، كلية الهندسة، جامعة بغداد.

٢) التميمي، أسامة عبد المنعم. (٢٠٠٦)، ***أثر تطورات تكنولوجيا المنشأ في التشكيل المعماري***، رسالة ماجستير، قسم العمارة، كلية الهندسة، جامعة بغداد.

٣) الثويني، د.علي. (٢٠٠٣)، ***مذاهب الهندسة المعمارية العراقية***، مقالات وآراء منشورة، المجلة المعمارية العراقية ISM، الشبكة الدولية.

[http://www.iraqisciencejournal.com/articles/200308/36]

٤) جعفر، علي محسن. (١٩٩٩)، ***الهيكل الإنشائي والمعنى في الشكل المعماري***، رسالة ماجستير، قسم العمارة، الجامعة التكنولوجية.

٥) حويش، عقيل نوري الملا. (١٩٨٨)، ***العمارة الحديثة في العراق، تحليل مقارن في هندسة العمارة والتخطيط***، دار الشؤون الثقافية العامة، الطبعة الأولى، بغداد.

٦) السهيري، عاطف. (٢٠٠١)، ***التكامل في الفعالية البنائية –النظرية وتطبيقها على البيت البغدادي التقليدي***، بحث منشور، مجلة اتحاد الجامعات العربية، العدد١، المجلد,٨

٧) السهيري، عاطف. (٢٠٠٥)، ***تكنولوجيا العمارة***: محاضرات أُلقيت على طلبة الماجستير في قسم العمارة، كلية الهندسة، جامعة بغداد.

٨) منتدى معماري. (٢٠٠٦)، ***العناصر الإنشائية وعناصر التشكيل***، مقال منشور، الشبكة الدولية.

[http://m3mare.com/vb/showthread.php?t=425]

9) Angerer, Fred. (1961)**,** ***Surface structure in buildings***, Alec Tiranti Ltd., London.
10) Geraldine Gildea. (1998), ***Structural Forms and Concepts***, B. Tech. (ED) Prog. University of Limerick press.
11) Graves, Michael. (1996), ***A Case for Figurative Architecture***: in Terrorizing a New Agenda for Architecture, Princeton Architecture Press, New York.
12) Holgate, Alan. (1986), ***The art in Structural Design***, Oxford University Press, London.
13) Nervi, P.L. (1956), ***Structures***, McGraw-Hill Inc., New York.
14) Salvadori and Heller**.** (1975), ***Structure in Architecture: The Building of Buildings***, Prentice-Hall Inc. Englewood Cliffs, New Jersey.

آلية عمل المبنى الصناعي
كنظام للمحافظة على الطاقة

الباحثون:

الدكتور بهجت رشاد شاهين
أستاذ مساعد
جامعة بغداد – كلية الهندسة
قسم الهندسة المعمارية

أمجد محمود عبد الله ألبدري
دكتوراه هندسة معمارية(مدرس)
جامعة بغداد – كلية الهندسة
قسم الهندسة المعمارية

Industrial building approach as an energy conservation system

ABSTRACT

Aiming at energy preservation and depending on the production of years of expertise and hard work of international research laboratories in the area of energy preservation and finding alternatives and undepleted energy sources in all aspects and various areas of the world, all these institutions asserted on following the Sustainable Design Method. It pursues the integrated building design method, which is one of the modern methods in the domain of energy preservation . This method focuses on the building's perfection from the very first stages of the design process, Through integration between the building components and mechanical services and supplementary systems in it and the impact exchange positively among them instead of being equipment used in a building that depends on the whole Building Approach , which has two key features , being circular and containing feedback concept , this makes it a frequent performing within the identification of the general conceptual framework.

Considering oil and other new sources as unreliable as continues energy sources and synchronically and interactionally with the products of knowledge developments and technical products aiming at diminishing energy consumption and producing the light and thermal environment comfortable for staff at the human environment generally and industrially environment specifically , the study directed its research problem toward clarifying and demonstrating substance and effectiveness the method of the integration industrial building on the internal human and mechanical environment, aiming at demonstrating its importance as a consistent work mechanism that participate in organizing and correcting courses of the internal productive environment , plus demonstrating effect of its performance in clarifying the sustainable design interacted with the external environment to preserve energy .

The aim is identifying the basic steps that role components of the integrated industrial building method to be a specific interactional method and a step context that explores work mechanism and the internal environment response to the

external one to turn it to a visually , biologically and productively comfortable work environment.

The research development on various studies and researches of many centers of the world , to be a base of the applied discussion of this study.

١.خلاصة البحث:

بقصد الحفاظ على الطاقة واعتماداً إلى ما تقدمت به وجالت بفيضه سنوات الخبرة والعمل المضني للمراكز والمختبرات البحثية العالمية في مجال الحفاظ على الطاقة وإيجاد البدائل لمصادر الطاقة غير الناضبة على كل الأصعدة وفي مختلف بقاع الأرض ، فقد أجزمت كل هذه المؤسسات على اعتماد منهج التصميم المستدام Sustainable Design الذي يعتمد أسلوب تصميم المبنى المتكامل والذي يعتبر من الأساليب الحديثة في مجال الحفاظ على الطاقة Energy .إذ يركز على مثالية المبنى منذ المراحل الأولى لعملية التصميم من خلال التكامل Integration بين عناصر أو مكونات المبنى (Building Components) والأنظمة الميكانيكية والخدمية والتكميلية الداخلة فيه والتأثير المتبادل بينهما بشكل ايجابي ، بدلاً من كونها معدات (Equipment) تستخدم في مبنى يعتمد على المنهج التكاملي (Whole Building Approach) ، الذي يمتاز بميزتين أساسيتين، في كونه حلقي وحاوٍ لمفهوم التغذية الاسترجاعية (Feedback)، مما يجعله تكراري الأداء ضمن تعريف الإطار ألمفاهيمي العام.

وعلى اعتبار أن النفط وبقية المصادر المستحثة لا يمكن أن يعول عليها كمصادر طاقة مستمرة . وتزامناً وتفاعلاً مع ما ينتج من تطورات معرفية وانتاجات تقنية تهدف إلى تقليل استهلاك الطاقة وإنتاج البيئة الضوئية والحرارية المريحة للعاملين في البيئة الإنسانية عموما والمصنعية خصوصا، فقد توجهت الدراسة بمشكلتها البحثية نحو توضيح وإظهار ماهية وفاعلية منهج المبنى الصناعي المتكامل على البيئة الإنسانية والميكانيكية الداخلية ،بهدف إبراز أهميته كآلية عمل تفاعلية تسهم في تنظيم سير وتصحيح مسارات عمل البيئة الإنتاجية الداخلية، مع إظهار اثر ادائيتها في توضيح فكرة التصميم المستدام المتفاعل مع البيئة الخارجية حفظا للطاقة. هادفاً بهذا نحو تحديد الخطوات الأساسية التي تحكم مكونة منهج المبنى الصناعي المتكامل، ليغدو أسلوبا تفاعليا محددا بخطوات يوضح آلية عمل واستجابة البيئة الداخلية إلى الخارجية، ليحيلها إلى بيئة عمل مريحة بصريا وبيولوجيا وإنتاجيا.

وقد اعتمد البحث في هذا على دراسات وبحوث متنوعة لمراكز متعددة في العالم لتكون أساسا في الطرح التطبيقي من هذه الدراسة.

٢. كلمات رئيسية:

الأتمتة ، حفظ الطاقة ، التصميم المستدام ، الايكولوجي ، نظام سريان المعلومات ، منهج المبنى الصناعي المتكامل، المماثلة الحية.

٣.المقدمة:

يقصد بالتكامل "Integration" "التشكيل الكلي أو التوحد مع شيء أخر للوصول إلى الكل" او يعرف التكامل بالتشكيل (To form) او المزج (To Blend) في الكل ، لمكونات فردية وجماعية مما يعطي الوحدة (Unity) . وتكون عملية التكامل هي الفعل (Act) او العملية (Process) في إيجاد هذه الوحدة ، وضمن أنظمة المبنى يكون مفهوم التكامل هو الفعل في خلق وظيفة المبنى الكلية الحاوية على أنظمة المبنى بكل متنوع ، ضامة لأنظمة الطاقة فيه من نظم (HAVC) ، ونظم الإنارة والتهوية ، ومع عناصر المبنى ، لتجهيزها بالطاقة اللازمة بالشكل الذي يعطي أدائية عالية للمبنى . (البعلبكي- ،١٩٨٠ص٤٧٢)، (Rush,1986,p4)،(Websters,1973 ,p600).

وهنا يعتمد منهج المبنى المتكامل على تحليل ادائية المبنى لكل ساعة ، باستخدامه موديل او نموذج افتراضي للمواقع في برنامج حاسوبي ، يزود التصاميم الاولية بمقدار الطاقة المخمنة، وليتم اجراء التعديلات التصميمية عليها باسلوب التغذية الاسترجاعية (Feedback) خلال عملية التصميم الاولية وصولاً للتصاميم النهائية النموذجية بدلاً من معرفة الطاقة المخمنة وكلفها في المراحل المتاخرة من التصميم، او لربما عند اكتماله ، هذا الاسلوب يساعد على تزويد المصممين بالمعلومات الدقيقة والمفيدة وتقودهم الى اتخاذ القرارات التصميمية الصحيحة والنموذجية، (عبيدات، ١٩٨٤،ص٨٩-١١٧) ، (BTS,2000,pp.1-4) ، (Snyder, 1979,pp.151-163)،(NREL, 2000 ,p.383).

لذا فان العملية التصميمية ضمن فكرة المبنى الصناعي المتكامل، تتصف بخصائص عامة ، اول ما يميزها به هو كونها حلقية، مع احتواءها على مفهوم التغذية الارجاعية (FeedBack) والذي يكمن في كونه سر حياة هذه العمليات ، اذ ان بعض المعلومات الجديدة قد تدفع المصمم الى اعادة النظر في المعلومات المتوفرة عنده كلما تطور التصميم ، مما يجعل العملية التصميمية عملية مكررة ، فالمصمم يمر خلالها بعدة مراحل في كل مرة يدخل عدد من المتغيرات في التصميم حتى يصبح التكوين اكثر نضجاً ، وان تكرار العملية لعدة مرات سيعمل على تحقق التصميم المطلوب. وثاني خصيصة لهذه الفكرة ، هي في كونها مقادة بمجموعة من المفاهيم والستراتيجيات التي تخط شكل العلاقات المتفاعلة وضمن اطار مفهوم التغذية الاسترجاعية ما بين المتغيرات المناخية والعوامل المؤثرة على انتاج شكل وقشرة المبنى، وما بين ادائية هذه القشرة وتقنياتها الضوئية ذات الاداء المتكامل لتجسيد فكرة المبنى الصناعي الكامل واهدافه في حفظ الطاقة والبيئة الداخلية المريحة للعامل.

٤. المشكلة البحثية والهدف:

وفقا لاعتبار تامين البيئة الداخلية المريحة للعمال داخل المصنع ، مع اعتبار تامين اعلى ادائية لعملهم والمكائن، وباقل قيم مصروفة من الطاقة لتامين مبدا خلق بيئة داخلية فعالة مؤتمتة مستجيبة اقتصادية، فقد حددت مشكلتي البحث كمايلي:

- مشكلة البحث العامة، التي انطوت تحت اظهار ماهية وفاعلية منهج المبنى الصناعي المتكامل على البيئة الانسانية والميكانيكية الداخلية ،بهدف ابراز اهميته كالية عمل تفاعلية تسهم في تنظيم سير وتصحيح مسارات عمل البيئة الانتاجية الداخلية، مع اظهار اثر ادائيتها في توضيح فكرة التصميم المستدام المتفاعل مع البيئة الخارجية حفظا للطاقة.
- اما مشكلة البحث الخاصة ، فقد اكدت اهمية اظهار سياق والية عمل وخطوات التشكيل والتداخل لمنهج المبنى المتكامل بصفته سياقا خطواتيا واسلوبا تفاعليا يحكم الية الحركة والتفاعل لعصب حياة المعمل المكائني الداخلي ليجعل منه كيانا مؤتمتا مستجيبا للنشاطات بصفة ذاتية.

وعليه يكمن هدف البحث – في تحديد الخطوات الواضحة التي تحكم مكونه منهج المبنى الصناعي بصفته برنامج عمل مغلق الدورة تحدد خطواته وعناصره مكونة الدورة الاسترجاعية Feedback ، لتجعل منه اسلوبا تفاعليا وسياقا خطواتيا يسير الية عمل واستجابة البيئة المكائنية والانسانية داخل بيئة الفضاء المصنعي ، ليحيلها الى بيئة عمل مريحة بصريا وبايولوجيا ، ليغدو هذا المنهج كلا متكاملا يحكم ويؤطر النظام الادنى ويحكمه ويؤطره كل اكبر ضمن منظومة اكبر يحكمها ويؤطرها النظام الاكبر الى ان تصل الى البيئة الام.

٥. منهج المبنى المتكامل:

ان عملية التحقق من مدى مصداقية هذا المنهج يتم من خلال مايسمى الحوافز المالية (Financial incentives) ، باستخدام مشروع حفظ الطاقة السنوي (Annualized energy saving project) للمباني الجديدة باعتماد اداة التحليل لصرف الطاقة ولترفع كمستوى لزيادة الكفاءة المخمنة ، والتي تستعمل على تخفيض كلف المشروع التقليدي المشابه بنسبة تتراوح مابين (٥٠-٨٥%) ، وذلك بتوظيفها للوسائل والتقنيات والاساليب الحديثة لقشرة وخدمات المبنى، بالاضافة الى اختيار الاجهزة والادوات والمكائن المتطورة الكفوءة التي جميعاً تمتاز بانها ذات استهلاك قليل للطاقة .

لذا فان اسلوب او منهج المبنى الكامل(Whole building approach) سيكون خادماً لكل من عمليتي التصميم والانشاء ، وذلك من خلال إتسامه بما يأتي (BTS, 2000,p212)، (NREL,1994.p362):-

١- استخدام الطاقة بكفاءة عالية تظمن تقليل استهلاكها .
٢- تحقيق الراحة الحرارية للفضاء الداخلي للمبنى باستعمال المعدات والاجهزة الميكانيكية الملائمة وتحسين الاجواء الداخلية صحياً وحرارياً ووظيفياً وصوتياً وبصرياً . مما يعطي تصور التصميم المثالي المريح بصرياً وبايولوجياً .
٣- توظيف واستخدام فوائد وتقنيات الضوء الطبيعي (Daylighting) بتكامل مع منظومة الاضاءة الصناعية .
٤- توظيف طاقة الاشعاع الشمسي الذاتية (Passive solar energy) ليعمل على تقليل كلف الطاقة العامة للمبنى بدون اضافة اي كلف اضافية فوق الكلف الرئيسية للمبنى .
٥- يمتلك هذا المنهج اهمية كبيرة في توظيف الوقت والجهد وتشغيل الاجهزة بشكل كفوء وبدون اضرار مع امكانية وسهولة الصيانة والسيطرة المؤتمتة المبرمجة.
٦- اضافة الى ما تقدم يعطي تكاملية تامة مابين التصميم والانظمة والفعاليات العملية ، مع اعطاءه انفتاحاً وكشفاً عن مكامن استخدام مصادر الطاقات المتجددة (Renewable energy sources). (SBIC,2001,p201). (NREL,1994.p373)، (Broadbent,1988,p71) ، (Littlefair,1997,p19).

٦- توازن منهج المبنى المتكامل:-

ان نطاق مفهوم الطاقة الذي يقصد التقليل من استهلاكه، هو تلك الطاقة المستهلكة كناتج نهائي ضمن حقل العمارة Architecture Field لتشغيل وادامة مستلزمات الراحة البصرية والبايولوجية في البيئة الداخلية، وان الحفاظ على الطاقة في المبنى يعبر عن التقليل في استهلاك المتجدد وغير المتجدد منها. حيث يقاس مفهوم المحافظة على الطاقة نسبة الى الحالة المثالية للاستهلاك المتضمن اقل مقدار لافضل الظروف والذي يكون لاسلوب تصميم المبنى مع البيئة المحيطة دور اساسي فيه لتقليل احمال المبنى من الطاقة الذي قد يحتاج صرفاً اضافياً منها للوصول الى حالة التوازن (لتوفير الاضاءة وتقليل الاحمال الحرارية – من تقليل الاكتساب الحراري الصيفي والفقدان الحراري الشتوي). وان الحفاظ يؤكد على التقليل الاكبر من استخدام الوسائل الميكانيكية العملاقة للحصول على متطلبات البيئة الداخلية للابنية الصناعية خاصة اذا علمنا ان كل وحدة من وحدات الطاقة الكهربائية يحتاج انتاجها صرف ثلاث وحدات من الطاقة الحرارية . وهنا ينصب دور المصمم والتصميم في الموازنة وتحسين العلاقة بين الاداء الحراري والمردود الاقتصادي (على المدى البعيد) ليصب في جدوى حفاظ الطاقة ، فهي عملية مستمرة على طول عمر المبنى والتي تعرف بـ(تحليل الكلفة لعمر المبنى) (life cycle costing analysis) وفقاً لتحقيق مبدأ شكل المبنى الصناعي المتوازن . (Julier, 2000,p12). حيث ان اتباع منهج التصميم المستدام Sustainable Design ، والسيطرة

المناخية على ادخال الضوء ودرء الاجهادات الحرارية للبيئة الخارجية ، سيكون هو المقيل في الوصول الى متطلبات الراحة الداخلية بصرياً وحرارياً . وهذا المنهج عادةً ما يكون متنوع الحلول فتارةً ياتي بشكل حلولٍ وقائية منذ بداية العملية التصميمية (شاملاً الموقع وكتلة البناء)، واخرى يكون بهيئة جرعات علاجية عند المراحل الحرجة من العملية التصميمية او حتى التنفيذية والتي عادة ما تتم بالتوظيف المقحم للوسائل الميكانيكية (Broadbent, 1988,p76)، (Mc Cluney, 1991,p108) .

وكثيراً ما كان الحلان يعملان معاً ، اذ انه ومن وجهة نظر العملية التكاملية، فان كلا الحلين يعتمد احدهما على الاخر ولا يمكن الفصل بينهما ، لهذا كرست جهود حثيثة لتصميم مباني ذات كفاءة طاقة مرشدة وبتوظيف الافكار التفصيلية والتخطيطية المطورة جامعةً العوامل السرمدية والعامة (كضوء الشمس وحركة الهواء والشكل باستجابة مناخية) ، وبهذا اصبح تقليلها حاجة اجتماعية وتقدم حضاري يحتم على المعمار ان يطور ويوظف حسه البيئي (ايكولوجي) لتحقيق اهداف التصميم المستدام في المحافظة على الطاقة وتوفير بيئة داخلية مريحة بصرياً وبايولوجياً . وان المحافظة على الطاقة سيعتمد على تحسين كفاءة الاجهزة والمكائن والمعدات الموظفة في المعمل وعلى تقليل الطاقة المصروفة لتشغيل وادامة هذه المعدات والاجهزة اضافة الى اجهزة ومعدات المبنى الخدمية اعتماداً على الخصائص التصميمية للمبنى وحجمه وفعاليته . (AIA ,1989,p310).

٧- <u>حفظ الطاقة بمنهج المبنى المتكامل:</u>

ولتقليل الطاقة فقد وضعت برامج عديدة من قبل مراكز بحثية متعددة في العالم:-

أ- التعامل مع الطاقة الحرة بتصميم مستدام يستغل طاقة الاشعاع الشمسي وذلك بالتوقيع الملائم والتوجيه والحجم الداخلي وشكل المبنى والتفاصيل الاخرى المدروسة ، وانتاج مشاريع وليدة الطاقة ، كحال الكثير من المشاريع الصناعية وغير الصناعية التي كانت نتائجها ان خفضت من كلف الاضاءة الصناعية الى (٧٥%) كونها تدخل الضوء الطبيعي المستثمر بتقنيات الغلاف الخارجي والتي عمدت الى ادخاله بصيغته الباردة Cool Light ، ولتخفض من كلف التبريد الى (٥٥%)، وقد كانت هذه النتائج لاختبارات اجريت في مناطق حارة مشمسة كالمملكة العربية السعودية والهند، لذا يجب على المباني ان تتعامل مع طاقة الاشعاع الشمسي بصيغ وتقنيات ذاتية وبمناهج متطورة،(Broadbent, 1988,p76).

ب- حفظ الطاقة يجب ان ينسجم مع متطلبات الانسان وسلوكه الشخصي الارادي واللاارادي الذي يعتمد التصميم المستدام كمعطى شكلي للعمارة، على ان يكون تقليل الطاقة لايعني ابداً تقليل الراحة .

ج- استغلال افكار وستراتيجيات الطاقة(بادخال الضوء الطبيعي والتظليل مع الكتلة الحرارية الخازنة) والمتداخلة مع نظم السيطرة على الطاقة وعلى الخواص المؤثرة على ادائية الطاقة للعناصر "الامتصاصية والانعكاسية" ونقل الطاقة والتشعع .

د- توفير موازنة طاقوية باختيار مواد الانشاء بمواصفات عالية في العزل الحراري المستخدم في انشاء الجدران والسقوف والتقنيات والزجاج المستخدم في النوافذ لتحسين نوعية الاضاءة من خلال تكامله مع الاضاءة الاصطناعية وتقليل نفاذية الاشعاع الشمسي، لتعمل كل من عناصر الانشاء الانفتاحية والصلدة بتكامل لتحقيق هذه الموازنة.

هـ- اعتماد غلاف المبنى (قشرته الهيكلية المغلفة الخارجية) كمنظومة بيئية متكاملة تؤمن الانفصال عن والاتصال مع البيئة الخارجية الطبيعية،ليعمل بصيغة جلد الكائن الحي في توفير وتامين البيئة الداخلية المريحة.(Girardet, 1998,p50)، (Lampert, 1999,p148)، (Julier, 2000,pp64-65).

ان مؤشر نجاح منهج التصميم الكلي المتكامل للمبنى الصناعي، هو ادراك ان كل انظمة المبنى تتصل وتتكامل بعلاقات مشتركة مترابطة ذات اعتماد متبادل ايجابي. ومن خلال التحليل النظامي لهذه المنظومات المتكاملة يمكن الحصول على كفاءة ادائية عالية للمبنى مع تقليل للكلف الاولية والتشغيلية المؤثرة فيه .(Lampert, 1999,p149)،(AIA,1989,p151).

٨- المبنى الصناعي كنظام للمحافظة على الطاقة:

لقد تناولت كثير من الدراسات موضوع المبنى كنظام يحافظ على الطاقة اذا ما تكيف مع الظروف الخارجية من خلال ادامته لنفسه كونه مجموعة مترابطة متداخلة من المنظومات المادية تنظم لغرض معين ، تعمل عندما تكون نتائجها مطابقة او مماثلة للاهداف او المقاصد ، وتكون مبنية على متطلبات معينة وضعت لاجلها . وان هذه المنظومات المادية هي التي تؤلف الاجزاء الرئيسية للمبنى (الغلاف الخارجي والهيكل الانشائي والفضاءات الداخلية والمكونات التكميلية التي تشمل وسائل وتقنيات الاضاءة والتكييف) . وهذا يعني ان لكل مجموعة من العناصر لها ميزة التفاعلية لزمان ومكان معينين ، والتي تهيكل نظامها وتمنحها سمات تميزها عن غيرها ، وتضم العلاقات الاتصالية والانفصالية مع بيئتها المحيطة . وبالتالي ستتصف طبيعة هذه النظم بصفتين ، هما صفة التنظيم وصفة العلاقات والتفاعل بين العناصر ، اللتان ستعطيان النظام مميزات تجعله من الممكن ان يستخدم في حل المشاكل بنفسه عن طريق منظومته الذكية المتكاملة، مما سيجعل الهيكل العام للنظم يتضمن مفهوماً اساسياً يعمل على الدوام على تدقيق النتائج مع الاهداف وهو التغذية الاسترجاعية (Feed Back) والتي هي اساس الربط بين مشابهة الكائنات الحية وعلاقتها بالبيئة (ايكولوجياً) مع مكائن الاحتراق الداخلي وكذلك الاساس في عمل نظرية التحكم الذاتي (Cybernetic theory) ونظرية النظم العامة "General system theory".(الضامن – ١٩٧٩-ص٢٨٢)،(Rush ,1996,p18)، (Handler, 1990,p22) .

١,٨- المماثلة الحية للمبنى الصناعي كنظام للمحافظة على الطاقة:

عند توظيف افكار "نظرية التحكم الذاتي" ، "نظرية النظم العامة"●، تصبح عملية المماثلة (Analogy) بين عمل الكائنات الحية والبنى يمكن ان تكون صحيحة ، فالمبنى بامكانه ان يمتلك صفة التنظيم كنظام مفتوح يديم استمراره من خلال التغذية الاسترجاعية المستمدة من متغيرات البيئة المحيطة عند تنظيم مفرداته ، والتي تكون ديناميكية ومتغيرة الاتجاهات تعمل على تزويد النظام بالطاقة وتعمل مع البيئة التي تحيط به بموجب نظام الموازنة الديناميكية

●

ان نظرية التحكم الذاتي ارتبطت بمفهوم الهوميوستاتيك(الاتزان البدني لعناصر الكائن الحي)(Homeostatic) حيث اوضح كانون "Canon" عام ١٩٣٩ ، الكيفية التي يتم بها بقاء درجة حرارة الجسم ثابتة واطلق عليها ما اسماه الهوميوستاتيك والتي لاتعني الجمود او بقاء الشيء راكد . اذ انها تعني الحالة التي يمكن ان تتغير ولكنها ثابتة نسبياً . ان الميكانيكية الهوميوستاتيكية للجسم تتطلب عدداً من المكونات المتصلة لكي تعمل مع بعضها بطريقة معينة معتمدة على ثلاثية وجود (عضو الاستقبال "مركز السيطرة" (الدماغ) المتحسس والفعل المناسب) و(التاكد من اتمام العمل بواسطة التغذية الاسترجاعية)، واصل كلمة (Cybernetic) هي من اصل يوناني (Kybernetes) والتي تعني "المتحكم" وقد استعملها مفهومها واط في ماكنته البخارية ليتمكن من التحكم بها ، لكن ما كان اكثر دقة هو ما توصل اليه " فينر " باطلاقه دورة التغذية الاسترجاعية والتي تعتبر الاساس في عمل ميكانيكية التحكم الذاتي والميكانيكية الهوميوستاتيكية للجسم المشتركة مع بعضها ، حيث يتم التحكم بها من خلال فكرة دورة التغذية الاسترجاعية (Feedback loop) ، ففكرة البحث عن الهدف في الميكانيكية الهوميوستاتيكية تختلف عن ترتيب عمل النظام لدى "واط" فماكنة "واط" لاتحاول السيطرة على البيئة المحيطة بها ، اما فينر فمحاولته بالسيطرة على الهدف المتحرك هي حالة اعلى مستوى ، اما نظرية "النظم العامة" (General system theory) فتعنى بالعلاقة والمشابهة بين فكرة النظم الحية ونظم المكائن "التحكم الذاتي" والعلاقة بين هذه النظم والبيئة ، فالكائنات الحية هي نظم مفتوحة تقيم علاقات تبادلية مع البيئة عكس المكائن فهي مغلقة امام البيئة . وقد وصف (برتا لانفي) (Bertalenffy) واضع نظرية النظم العامة ، بان النظم المغلقة لاتدخل او تخرج منها مواد وتكون "مفتوحة" عندما يكون لها خروج ودخول ، ولهذا تتغير مكوناتها فالنظم الحية تحمل هذه الصفة الاخيرة وتديم نفسها من خلال التبادل بالمواد مع البيئة وبحالة مستمرة تبني وتهدم مكوناتها بصورة تفاعلية. (Montgomery, 1998,p20)،(سابرينا– ١٩٧٦ ص١٧٦– ١٧٧) .

بينهما (Dynamic Equilibrium)، والتي يكون فيها غلاف المبنى النفعي اساس الانفصال والاتصال (Handler, 1990,p25) .

وهنا نستطيع اعتبار المبنى نظاماً يمتلك الصفات التي تمكنه من ادامة عملياته بفاعلية اذا تم اعتباره تكويناً مادياً حياً قادراً على ان ينطلق الى جميع اجزائه ويتفاعل معها ومع البيئة التي تحيط به عبر مكونات غلافه ، وهو اهم ما يميز النظم الحية ، حيث ان حيوية المبنى وقشرته هو ان يقترب من حيوية الكائن الحي (بجلده) في ادائه ، طالما كان التكامل لاعضاء الكائن الحي الداخلية والخارجية اساساً في الاتصال والانفصال ، حيث ان خصوصية المبنى التي اوجدت باحتياجات وظيفة العمل ومتطلباتها ستكون متكاملة رغم عدم اعتبار تلك الخصوصية مؤهلة للتعبير عن انظمة خاصة ، على اعتبار ان هذه الخصوصية متكاملة بالتعريف عن الظروف الخارجية بردود افعال الغلاف وعن الظروف الداخلية بتفاعل انظمة الغلاف والبيئة الداخلية (اعتماداً الى الوظيفة والزمان والمكان والشاغلين)، لكنها دائمة التغير والتفاعل لتضفي صفة الديمومة المستدامة . وبهذا يمكن اعتبار المباني الموظفة لفكرة المبنى الصناعي المتكامل بقشرته الفعالة، مبانٍ ذات تكامل دائم التغير لعدم استقرار- وسكون مفاصل- تلك المباني ليعطيها صفة التواصل الزمني والمكاني (وهو مقياس الاستدامة) .

ان مدى ارتباط النظام (المبنى) بالمحيط الخارجي يمكن من خلال مقياس الانتروبي (الذي هو الطاقة الفائضة غير المستفاد منها في نظام الثرموداينامكي المغلق)(Closed thermodynamic system) ، اذ ان الكائنات الحية تقتات على "الانتروبي" السالب وان النظام الحي يديم نفسه بحالة توازن من خلال استيراد المواد الغنية بالطاقة عكس النظم المغلقة ، لتعمل ردود افعالها الداخلية على زيادة "الانتروبي"، مما يعطيها التوازن لحصول تحولات واضحة محددة مابين مصادر الحرارة والعمل المؤدى. وحيث ان الطاقة عند تحويلها من شكل لاخر لاتتحول جميعها الى شكل اخر للطاقة، فان بعضها سيبقى بشكل لايمكن الاستفادة منه هو (الانتروبيا)، مع ان شيئاً لايفقد منها لانها تبقى ثابتة داخل النظام . وهذا ما يجعل النظام الكوني يميل للحفاظ على الطاقة من خلال محاولة بذل اقل طاقة لكل فعالية او عمل مطلوب، الامر الذي ينطبق على الكائن الحي والمادة غير الحية ، فالماكنة لاتعمل بكفاءة مثالية (١٠٠%) مهما بلغت من الجودة والدقة في الصنع. وعلى هذا الاساس فان هذا يستلزم من نظام المبنى ان يبقى محافظاً على الطاقة لتوفير بيئة داخلية مريحة وثابتة نسبياً، يبذل فيها الانسان اقل طاقة او مجهود لادامة مفهوم الراحة البصرية والبايولوجية الداخلية اللتان ستنعكسان بالتالي عليه لتوفير مضاعف لمفهوم الحفاظ على الطاقة بشكل ديالكتيكي.(Lampert,1999,p301)، (Johnson, 1991,p542)،(سابرينا- ١٩٧٦ ص٢٦٩).

٢,٨- <u>الاتمتة المسيطرة للمبنى الصناعي كنظام للمحافظة على الطاقة:</u>

وبموجب نظرية التحكم الذاتي ونظرية النظم العامة فان عملية المماثلة الحية اظهرت سياسات الاتمتة والسيطرة في المعامل الصناعية وميكانيكياتها ، تحكم كلتا العمليتين التصنيعية والبيئية المجهزة بواسطة التقنيات والوسائل التي تنفذ الاوامر الصادرة من النظام العام . وخلال السنين المنصرمة ، فان الاتمتة Automation لهذين الحقلين المختصين بالسيطرة قد طورت الى تطبيقات عمل وقواعد تكنولوجية مستقلة منفصلة موجهة لحقل الصناعة على وجه الخصوص، اذ ان هنالك اسباباً اكراهية وموجبات الزامية اقتصادية لتداخل وتكامل السيطرات التصنيعية والبيئية ونظمها التقنية من اجل بلوغ اهداف حفظ الطاقة واعطاء بيئة داخلية مسيطر عليها مفعمة بالحياة والنشاط .

وقد أوجدت هذه القواعد التكنلوجية التطبيقية في مجال الاتمتة والسيطرة على البيئة الصناعية ان هنالك ثلاث ابعاد من تداخل النظام (سريان المعلومات)، هي:-

البعد الافقي – بين المعلومات ذات التداخل للسيطرات البيئية وتداخل سيطرات الانتاج الداخلية (تصميم تفاصيل الغلاف الخارجي ، تصميم منظومة الكهرباء الداخلية) .
البعد العمودي – بين المعلومات الاجرائية Transactional ونظم السيطرة التنظيمية بين الوظائف الصناعية (التشغيل للمكائن) .
البعد المحوري – طولياً على امتداد عامل الوقت ، لتنظيم وعي الانسان من خلال تنسيق عمل تقنيات الغلاف الخارجي مع البيئة الداخلية والخارجية .

ان كل ابعاد التداخل هي مهمة لتحسين وتحفيز نوعية الانتاج واظهار الاهتمام بعامل الاقتصاد وحفظ الطاقة.(شكل ١)،(Alarcon,1992,pp12-14)،(Blair, 1983,p83).

ان نظم السيطرات البيئية المقصودة عند البعد الافقي تضم كل مايخص خدمات المبنى الصناعي التحتية (كتقنيات السيطرة على عوامل المناخ والاضاءة والامان والطاقة). وقد اوصت منظمة (ASHRAE) الامريكية بان ادائية الانسان داخل المعمل ترتبط بصورة مباشرة بمديات خاصة محدودة (مريحة) من درجات الحرارة والرطوبة والضغط والاضاءة ونوعية الهواء وكذلك العوامل النفسية مما يتعلق منها خصوصاً بالارتباط مع المناظر الخارجية (وكل حالة بحالتها) ، وهي من عوامل تحديد امان العمال وصحتهم والابقاء على حيويتهم وتحفيزهم لزيادة الانتاج ، حيث يجب ان تلعب خصوصية البيئة الداخلية وفي مختلف حقول الصناعة دوراً مهماً في قابلية التطبيق والنجاح للجوانب الانسانية والتشغيلية المتعلقة بالمنتج والانتاج والامان والاقتصاد والصيانة واستهلاك الطاقة وتقليل الكلف. (BRT, 1999,p52)، (Fergusson, 1993,p93).

ان نظم السيطرة وعلى مختلف الاصعدة التقنية والتنظمية في المعامل لها سلسلة معقدة جداً ومستويات مختلفة في السيطرة واعطاءها للاوامر بنمطية تصاعدية Hierarchy للسيطرة على نظم التفاعل والاستجابة المناخية والانتاجية . وبموجب النظرة العالمية الجديدة والتطور الفني الذي تفرضه نظريتي النظم العامة والسايبرناتك (Cybernetic & General system theories) ، فان نظم مباني المعامل على اختلاف حجومها وانواعها ، تسكن نظم الاتمتة المتسلسلة بصورة تصاعدية Hierarchy بستة مستويات (6 levels) متهيكلة، تعرف مستويات الاتمتة الابتدائية والتي تتعقد وتكبر اكثر فاكثر كلما زاد التفصيل وارتفع مستوى التواصل مع الامر الاعلى (النظام المصدر الامر).(شكل ٢)،(Fergusson, 1997, P8)، (Blair, 1983,p55).

بموجب قوانين السايبرناتك فان نظم الاتمتة في المعامل تتركز اهميتها بالسيطرة والتحكم على تقنيات الغلاف والسيطرة على البيئة الداخلية والانتاجية مع السيطرة على حركة العمال والمكائن ، وهذه النظم تحكم من قبل مجموعة من بديهيات ومبادئ النظام والتي هي:-

<u>بديهية ١</u>- النظم على مختلف المستويات تحكم من قبل النظام الاعلى مرتبة .
<u>بديهية ٢</u>- ان النظام الاعلى يحكم من النظام الاعلى منه عند الانتقال بالمستوى الى مستوى اعلى منه .
<u>بديهية ٣</u>- التداخل لنظم السيطرة والاتمتة والخدمات قد يحدث افقياً و/او عمودياً و/او محورياً (زمنياً) او بمجموعها.
<u>بديهية ٤</u>- السيطرات المتداخل تعتمد على الاتصالات المتراكبة القوية مع بروتوكولات معرفة بصورة جيدة (مبرمجة) بين وعلى طول خط العناصر والتقنيات لمستوى معين وكذلك العناصر والتقنيات للمستويات المجاورة بنفس المستوى .
<u>بديهية ٥</u>- ان النظام المطبق (الثابت بصورة ديناميكية) يجب ان يكون مرناً وقابل للتبني adaptation ، لان البيئة التي تشغله متغيرة باستمرار .

بديهية ٦- ان النظام المطبق يجب ان يكون ذا تغذية مرجعية مرتدة على نفسه Reflexive معرفة حالاته وسلوكياته في سبل توفير امكانية التشخيص diagnosis للحالات الطارئة والصيانة المتوقعة مع التصليح الذاتي واعادة تنظيم تقنيات ووسائل السيطرة بموجب الحالة الجديدة .

بديهية ٧- ان النظم تتداخل فقط عندما تتقاسم لغة عامة (رمزية syntax) ومتفقة في المعاني والمقصود meanings عند الرسائل التي تغيرها داخلياً والتي تعمل كمختارات للسلوك selectors ومتغيرات للحركات وردود الفعل تجاه تغير حالة المناخ او البيئة الداخلية للمصنع .(Fischer, 2001,pp333-346)، (Katz, 2002,pp81-104)،(Keidel, 1985,p95) .

وبموجب هذا، فان هذه المجموعة من البديهيات ستعمل وفقاً لست مستويات في النظام التصنيعي هي:-

- المستوى الخامس (المؤسسة) - يضم الادارة وعمليات التطوير والتخطيط وخط انتاج الافكار وهو اعلى مستوى للاوامر ، ويمتاز بمسؤولية الاشراف على المستويات الادنى من السلوك .
- المستوى الرابع (المصنع) - الاشراف على الافرع وعلى نظم الانتاج ووسائل السيطرة على البيئة وهو يحكم المستوى الادنى منه.
- المستوى الثالث (المبنى/ الانتاج) - يسيطر ويدير ويشرف على كل خطوط المبنى والانتاج وينظم بيئتها وتقنياتها الداخلية ، وهو يشرف على المستوى الادنى منه .
- المستوى الثاني (المساحة/الخط) - يسيطر ويدير ويشرف على الادارة والسلوك لمساحة معينة من خط الانتاج، وهو يسيطر على المستوى الادنى منه .
- المستوى الاول (الوحدة/ الخلية) - يسيطر ويدير ويشرف على نظم التحكم بالحالات والسلوك لوحدة الاتمتة وخلية التصنيع ، ويسيطر على المستوى صفر(٠) ويجهز الاشراف على وسائل هذا المستوى (الحساسات والمشغلات الميكانيكية).
- المستوى صفر (الوسيلة) - يشمل الوسائل والتقنيات المتحسسة (sensetive)، التي تؤمن التشغيل الميكانيكي للعمليات الفيزيائية داخل نظم الانتاج والبيئة.(Howard, 1998,pp31-32)،(Matthews, 1989,pp14-16)،(Nam, 2002,p117) .

١,٢,٨- <u>الية نظام السيطرة:</u>

هنالك العديد من التشابهات والاختلافات بين الاتمتة للعمل ولنظم السيطرات البيئية والصناعية ، اعتماداً على طبيعة التصميم الوظيفي والمتطلبات ، والتي بصورة عامة تحكم الاداء والعمل المستمر المتحسس والمستجيب لتقنيات غلاف المبنى الصناعي البيئية ، وكذلك نظم السيطرة على البيئة الداخلية الطبيعية الانسانية والميكانيكية .

ان مفتاح التمييز بين نظم العمل والسيطرة قد اوجد في رياضيات وعلوم السايبرناتك من خلال الاتصالات والسيطرة الموجودة في النظم المتفاعلة interaction ، اذ ان الفكرة المركزية في هذه العلوم هي السيطرة المنظمة بالاعتماد على سيطرة الاسترجاع FeedBack control ودورات السيطرة بواسطة اداة التحكم المؤازر Servo control(شكل ٣)، والتي ستعمل على تنظيم عمل تقنيات السيطرة الخارجية والداخلية (البيئية والانتاجية) المسيطر عليها بواسطة المتحسسات لملاحظة سلوكها وحركاتها ، ومن ثم ستصدر معلومات قياسية باتجاه المستوى الاعلى من خلال الادراك لتحديد وتخمين حالات العملية والسلوك للمرحلة والحالة القادمة المتوقعة وغير المتوقعة.(Fischer, 2001,p212)،(الضامن- ١٩٧٩ص١١٧).

هذه القياسات الجديدة ستستعمل لتحديث update قاعدة المعرفة Knowledge Base لعملية السيطرة المبرمجة العامة للمعمل ، ومن قاعدة المعرفة هذه فان مبرمج السيطرة العام (الموديل) سيحسب ويتوقع الحالات القادمة ليتكيف معها ويجعلها قابلة للسيطرة عند العملية من

جديد مرسلاً الاوامر لواحدة او اكثر من المشغلات Actuators لتاثير واحداث التغيير في حركة وموقع واستجابة تقنية السيطرة البيئية او الانتاجية ، مما يجعلها وسيلة سيطرة منظمة لادائية وعمل التقنيات بنجاح لتوفير بيئة داخلية مريحة ومقنعة لشاغليها.(Keidel, 1985,p102)،(Fergusson, 1997,p25).

ولتطوير موديل النظام المتداخل ، فانه من الضروري البدء بتحليل دورة السيطرة Decomposing loop (شكل ٣) الى مكوناتها الاساسية Core Components عند مستوى السيطرة صفر المسيطر على العملية (L0) بفصل وعزل العناصر التكوينية الصلبة hardware والمرنة Software ، لتصبح التقنية وتوابعها من المتحسسات والمشغلات ضمن نظام hardware، اما ماينقل اليها وعنها من معلومات للسيطرة من قبل الحكم المنفذ بالحاسبة العليا المسيطرة ، فهي ضمن نظام Software ، مما يسهل ايصال المعلومات والسيطرة على مستوى (L0) بواسطة الحقل الرقمي العام (Digital field) . وان نفس وسائل المستوى الاوطأ (L0) ستحتوي المعلومات البيئية والانتاجية وحاجات العامل والصيانة وقاعدة المعلومات العامة للتشكيل المتجدد حسب الظروف والحالات. وهذا التوجه سيجعل من وسائل الحقل التطبيقي متزايدة الذكاء مما يسهل التحكم والحكم والتفاعل عند دوراتها الاسترجاعية (FeedBack)،(شكل ٤). (Stinchcombe, 1995,185)،(Graedel, 2003,pp342-345) .

وعليه فان المنطق المسيطر على الوسيلة (التقنية) للمستوى صفر (L0) هو انها ستكون موضوعة بصفة خادم مسيطر منه وعليه في التسلسل المتصاعد hierarchy باتجاه المسيطر الاعلى ضمن المستوى الاعلى .هذا التشكيل سيتطلب شبكة عمل متداخلة محسوبة للحالات الطارئة بوجود حساسات ومشغلات الية سريعة الاستجابة لانجاح مثل هكذا تفكيكات (شكل ٥).(Stinchcombe, 1995,p191).

٢,٢,٨- <u>توزيع اوامر نظام السيطرة:</u>

ان تصميم نظام السيطرة يثبت (بديهة ١) بان نظم المستوى الاوطأ محكومة ومضبوطة بمنطق المستوى الاعلى وان هذه العلاقة المنظمة هي حرجة عملياً كونها تداخل وتكامل ادارة عمليات مختلف المستويات لما فوق وتحت عصا السيطرة . حيث ان المسيطر التكاملي Unitary في المستوى (L1) يشرف على المسيطر للوسيلة (التقنية) عند المستوى صفر (L0). والحالة الراهنة عند المستوى (L0) تغذي باتجاه الاعلى حيثما مسيطر المستوى الاول (L1) يمكن ان يوظف معلومات للسيطرة على استراتيجية السيطرة عند المستوى (L0) . هذه العلاقة موجودة بين اي مستويين متراكبة وايضا بين المسيطرات عند نفس المستوى.

كذلك فان المسيطر المستوى الاول (L1) يلاحظ الحالة والسلوك الحركي والسيطرة عند عمليتي المستوى صفر (L0)، لضمان التساوق والتعاون في خط العمليات المتظمنة (الثبات Sable، والتزامن في الاداء Synchronized، والسماح بالعيوب البسيطة غير المؤثرة على الاداء العامFault tolerant حتى يبقى عاملاً)، (شكل٥).(Tatum, 1990,pp112)، (Nam, 2002,p133).

٣,٢,٨- <u>الدورة الاسترجاعية لنظام السيطرة:</u>

ان اساس الحياة والعمل عند كل مرحلة ومستوى من المستويات ابتداءاً من المستوى صفر (L0) والى اعلاه وعلى مختلف تفاصيلها ووسائلها تكون جميعاً محكومة ومفعلة بموجب نظام التغذية الاسترجاعية (FeedBack) . حيث ان العمليات على اختلاف مستوياتها مكونة من اكثر من عملية ثانوية محتواة interwinded (كالتخطيط الاستراتيجي ، الانشاء ، هندسة وتصنيع المكائن ، التنظيم ، ....الخ)، هذه العمليات الثانوية مكونة بذاتها من عمليات ثانوية

اخرى وهكذا . وان لكل من هذه العمليات على اختلاف مستوياتها اربع خطوات اساسية هي **المخطط Plan**-(لملائمة اهداف العمل الاستراتيجي) ، **العمل Do**-(المطلوب تنفيذه وهنا دور الوسائل والتقنيات)،**التاكد Check**-(من العمل المؤدى لتحديد اذا ما كانت الاهداف الستراتيجية الموضوعة محققة) ،واخيرا **النشاط Act**-(بواسطة تنظيم العملية لتحقيق النتائج الافضل). وهذه العمليات جميعاً تمتلك دورات استرجاعية تفصيلية داخلية ثانوية ودورات استرجاعية علائقية فيما بينها مع وجود الدورات الاسترجاعية العلوية التي تحكم علاقتها بالنظم والمستويات الاعلى (Tatum, 1990,p75)،(Thomas, 1993,p88)،(Vanegas, 1998,p125).

٩- الاستنتاجات :

وفقا لمنظور منهج المبنى الصناعي المتكامل ستكون المباني التي تصمم اعتباراً لسماته وخصائصه:

١- مبان عالية الأدائية (High-performance buildings) غير مستهلكة للطاقة حافظًة لها (Energy-saving building) مقابل ذلك تحقق متطلبات البيئة الداخلية من إضاءة وحرارة ونقاوة الهواء وتقليل الكلف الفعلية للمبنى، كونها تستعمل مقاييس كفاءة الطاقة المتجددة مما يطيل من عمر المواد المستخدمة ،

٢- ذات تصميم حساس بالموقع لتقليل الأحمال البيئية للموقع والمناخ، مع توفر عامل المرونة والتوسع المستقبلي وسهولة الصيانة والتحفيز لنشاط العاملين داخل المصنع، من خلال وضع العلاقة الحية الرابطة بين البيئتين الخارجية والداخلية بعنصر الوسيط (الغلاف) المتحرك بفاعلية تفاصيله وتقنياته.

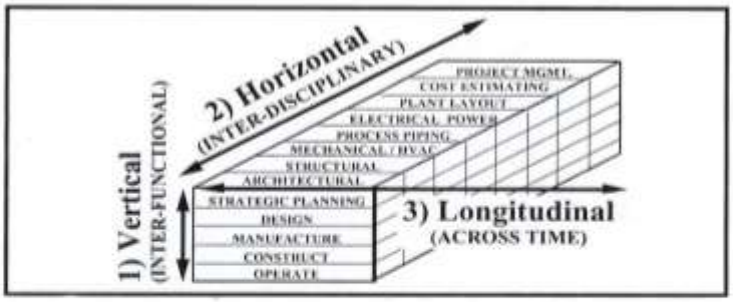
٣- ان تصميم المبنى وفق منهج المبنى الصناعي المتكامل، يدرس العلاقة بين حجم الفتحات مقابل حجم النظام الميكانيكي والتي يترتب عليها كمية الاضاءة الطبيعية، فضلا عن ساعات العمل ، طريقة تنظيم الفضاء وكفاءة الغلاف الخارجي للمبنى ، تاثير الموقع والمناخ الموضعي (local microclimate) والاتجاه والادائية العالية المتعددة منها (الوظيفة والحرارية والصوتية والبيئية والبصرية) ، كذلك فان هذا المنهج لايساعد فقط في تقليل كلف الانشاء وانما يقلل كلف التشغيل على طول الوقت.

٤- ليكون هدف التصميم المناخي المستدام للمبنى الصناعي من وجهة نظر المحافظة على الطاقة بانه عملية الحصول على اكبر تعرض للبيئة الخارجية ليوفر اقل استهلاك للطاقة حفاظاً وتثبيتاً لمفهوم الراحة الداخلية والتي يكون فيها كل الجهد المبذول من قبل العاملين داخل المبنى يصرف من اجل الانتاج الفيزياوي العقلي دونما فقدان او هدر يذكر.

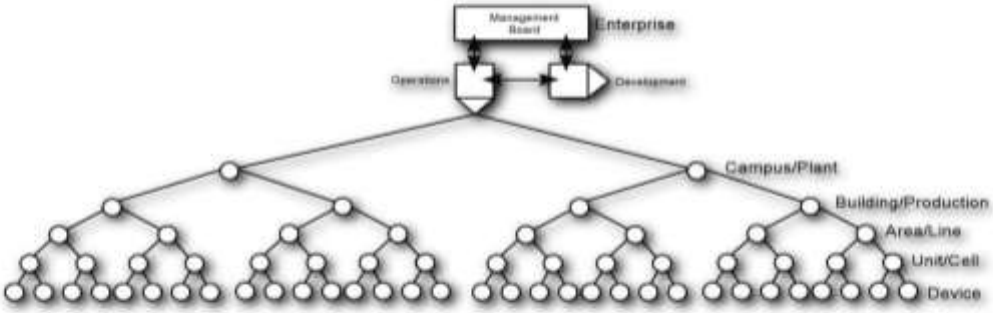
٥- ان نظم السيطرة وعلى مختلف الاصعدة التقنية والتنظمية في المعامل لها سلسلة معقدة جداً ومستويات مختلفة في السيطرة واعطاءها للاوامر بنمطية تصاعدية Hierarchy للسيطرة على نظم التفاعل والاستجابة المناخية والانتاجية.

٦- ان هذه القياسات الجديدة ستستعمل لتحديث update قاعدة المعرفة Knowledge Base لعملية السيطرة المبرمجة العامة للمعمل ، ومن قاعدة المعرفة هذه فان مبرمج السيطرة العام (الموديل) سيحسب ويتوقع الحالات القادمة ليتكيف معها ويجعلها قابلة للسيطرة عند العملية من جديد مرسلاً الاوامر لواحدة او اكثر من المشغلات Actuators لتاثير واحداث التغيير في حركة وموقع واستجابة تقنية السيطرة البيئية او الانتاجية ، مما يجعلها وسيلة سيطرة منظمة لادائية وعمل التقنيات بنجاح لتوفير بيئة داخلية مريحة ومقنعة لشاغليها.

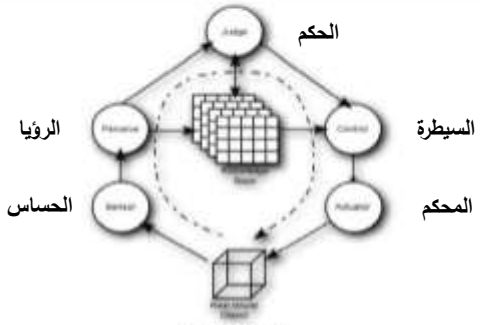
وعلى ضوء ما تقدم يمكن القول ان نظم السيطرات البيئية والانتاجية المتكاملة والمسيطرة في المعمل ستتماكن عند الخط الرئيسي للتطبيقات العملية الممثلة للقاعدة الوظيفية للنظم المؤتمتة المتكاملة والمتداخلة لتأدية عمل جسم المصنع البيئية والانتاجية. حيث يمثل هذا الخط التطبيقي مجموعة من الحلقات المتراكبة تمثل المستويات الستة للاتمتة المبينة سابقاً ، والتي تتمثل بسيطرات التصنيع الانتاجية و البيئية ، متكاملة بعملها مع منظومة الادارة المالية والتخمين ونظم الانتاج المفتاحية.



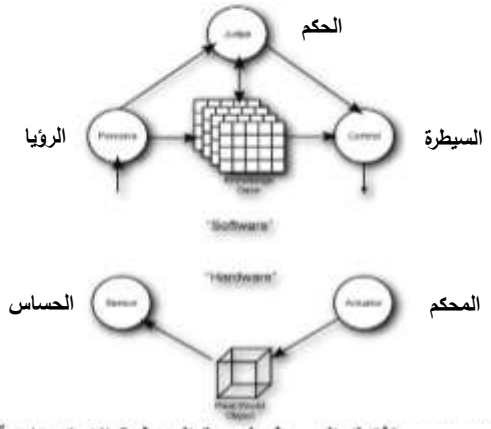
شكل (1) الأبعاد الثلاثة لتداخل نظم الأتمتة وسريان المعلومات المسيطرة على المبنى الصناعي الكامل المحافظ عل الطاقة (Alarcon,1992,p4)



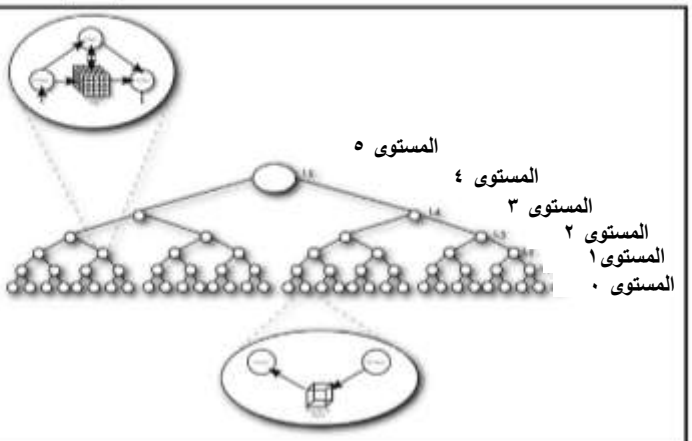
شكل (2) هيكل السيطرة و الأوامر لنظم الأتمتة المتسلسلة بصورة تصاعدية ذات ستة مستويات Ferguson,1997,P8



شكل (3)
دورة السيطرة الإسترجاعية
(Fischer, 2001,p212)



شكل (4) تفكيك المسيطر لدورة السيطرة الإسترجاعية
(Stinchcombe, 1995,185)



شكل (5)
لسيطرة الموزعة لدورة السيطرة الإسترجاعية
(Stinchcombe, 1995,p191)

١٠- المصادر الأجنبية :

- AIA ,1989, " Window Configuration " – Designing for Daylighting and Productivity, The Ehrenkrantz Group.
- Alarcon –Caedenas,L.F., /Ashley , D. B., 1992, " Project Performance Modeling: Methodology for Evaluating Project Execution Strategies ", Report to construction Industry Institute.
- Blair, R. D.,/ Kaserman, D. L., 1983, " Law and Economics of Vertical Integration and Control ", Academic Press, Inc., New York .
- Broadbent, Geoffrey, 1988," Design in Architecture ", Library of Congress, Lowcain, Belgium.
- BRT, Business Roundtable, 1999, " More Construction for the Money: Summary Report of the Construction Industry's Cost Effectiveness Report ", Business Roundtable, New York .
- BTS, 2000, " Whole Building design ", U.S. Department of Energy, office of building, London .
- Fergusson, K.J./ Teicholz, P. M., 1993, " Impact of Integration on Industrial Facility Quality ", Department of civil Engineering dissertation, Stanford University .
- Fergusson, K. J./ Teicholz, P. M., 1997, " Owner Perspectives on Industrial Facility Quality ", Journal of Performance of Constructed Facilites,, Vol.18, No.2, ASCE .
- Fischer, M., 2001" Design Construction Integration Through Constructibility Design Rules for the Preliminary Design of Reinforced Concrete Structures ", presented at the CSCE/CPCA Structural Concrete Conference in Montreal, Canada .
- Girardet, Herbert, 1998, " The Architecture of Ecology ", Academy Editions Press, London.
- Graedel, T.E./ Y. Kakizawa,/ M. Jensen., 2003, " Industrial Ecology and Automotive Systems ", A Handbook of Industrial Ecology , Edward Elgar, Northampton .
- Handler, A, Benjumin; 1990, " System Approach to Architecture "; American Elsevier publishing company Inc; New York .
- Howard, H. C./ Levitt, R. E./ Paulson, B.E./ Pohl, J. G.,/ Tatum, C.B., 1998, "Computer Integration: Reducing Fragmentation in the AEC Industry", Journal of Computing in Civil Engineering, Vol.3, No.1.
- Johnson, T., 1991, " Low-emissive Glazing Design Guide ", Butterworth-Heinemann. Stoneham, U.S.A.
- Julier, Guy, 2000, " The culture of Design " sage press, London.
- Katz, R., 2002 , " The Effects of Group Longevity on Project Communication and Performance ", Administrativ Science Quartrly, Vol.127.
- Keidel, R. W., 1985, " Game Plans ", Dutton, New York .
- Lampert, C., 1999, " Chromogenic Swithable Glazing: Towards the Development of the Smart – Window ", in the Proceedings of Window Innovations Conference, Toronto .
- Littlefair, Paul. J., 1997, " Designs for Improving Daylighting " Building Research Establishment (BRE), Watford, England.
- Matthews, M. F./ Burati, J. L. Jr., 1989 , "Quality Management Organizations and Techniques ", report to the Construction Industry Institute form Clemson University .
- Mc Cluney, Ross, 1991 " The Importance of the IDMY ", the (CIE) International Daylight Measurment Year, New York .
- Montgomery, Richard H, Mdes, Watter f; 1998, "The Solar Decision Book of Homes", John Wiley & Sons Irc; Conada.
- Nam, C. H.,/ Tatum, C. B., 2002, " Non Contractual Methods of Integration on Construction Projects ", Journal of Construction Engineering and Management, Vol.118, No.2, ASCE .

- NREL, 1994, " DOE's Passive Solar Nonresidential Experimental Building Program " Produced for the U.S.A. Department of Energy, by The National Renewable Energy Laboratory. New York .
- NREL, 2000, " Savings by design", Southern California Edison, Energy design resources, Vol.112 , No.1002, U.S.A,.
- Rush, Richard, 1986 " The Building Systems Integration Handbook ", John Wiley & sons. Ins ;New York; U.S.A;.
- Rush R.D:, 1996, " The Building Integration Hand book"; the American institute of Architests; New York .
- SBIC, ,2001, "The Whole Building Design Approach", Sustainable building industry council,; U.S.A.
- Snyder , J., 1979, " Introduction to Architecture Design and the design process ", Mc, Graw–hill Book Co. .
- Stinchcombe, A. L.,/ Heimer, C. A., 1995, " Contracts as Hierarchical Documents ", Organization Theory and Project Management, Norwegian University Press .
- Tatum, C. B., 1990, " Inegration: Emerging Management Challenge ", ASCE Journal of Management in Engineering, Vol.6, No.1, January .
- Thomas, V.C., 1993, " Buiding Systems Integration " Proceedings of the International Symposium on Building Systems Automation-Intrgration, University of Wisconsin-Madison .
- Vanegas Pabon, J., 1998, " A model for design/construction integration during the initial phases of design for building construction projects ", Department of Civil Engineering dissertation, Stanford University .
- Websters, 1973, " New Collegiate Dictionary ", G&c. Merrsiamco ,USA.

١١- المصادر العربية:

- البدري ، امجد محمود عبد الله- ٢٠٠٦- التطور والتغير في الفكر الجديد لعمارة الأبنية الصناعية الذكية- أطروحة دكتوراه - جامعة بغداد .
- البعلبكي ، منير- ١٩٨٠- قاموس المورد (انكليزي- عربي)- دار العلم للملايين- الطبعة الرابعة عشر- بيروت .
- الضامن، د.حاتم صالح- ١٩٧٩- نظرية النظم- تاريخ وتطور- الموسوعة الصغيرة (٤٧)- منشورات وزارة الثقافة والاعلام- بغداد .
- سابرينا ، د.آلن- ١٩٧٦- "بين الانسان والآلة - السايبرناطيقيا في داخلنا "- ترجمة صبحي ابو سعد- دار الكتاب العربي للطباعة والنشر- القاهرة .
- عبيدات، ذوقان/ كايد، عبد الحق- ١٩٨٤-" البحث العلمي، مفهومه واساليبه وادواته"- دار الفكر للنشر والتوزيع- عمان .

EFFECT OF PRESTRESSING FORCE ON TORSION RESISTANCE OF CONCRETE BEAMS

Prof .Dr .Husain M. Husain
University of Tikrit

Assist. Prof. Dr .Nazar K. Oukaili
University of Baghdad

Muyasser M. Jomaa'h
University of Tikrit

ABSTRACT

This study presents an experimental and theoretical investigation of torsion behavior of prestressed concrete rectangular beams without ordinary (or typical) reinforcement. Two concrete beams with concentric prestressing tendons (6-strands of 7 wires) and two plain concrete beams were tested in this investigation with f'c= 44MPa was used.

Experimental results showed that the ultimate torsional strengths increased by about 70% for the tested beams containing concentric prestressed strands over the plain concrete beams. Also the angle of twist decreased (68.8%).Crack patterns and the effect of compressive force due to prestressing tendons and high strength concrete can be denoted from the helical mode of single crack at midspan of the beams under testing and from the sudden failure mode. In the analytical work P3DNFEA (Program, three-dimensional nonlinear finite element analysis), by Al-Shaarbaf has been utilized. Three dimensional nonlinear quadratic 20 -node brick elements were used to model the concrete, while, the prestressing strands were modeled by embedded representation. Reinforcing bars (Prestressing strands) were assumed to be capable of transmitting axial forces only.

It was found that the general behavior of the finite element showed good agreement with observations and results from the experimental tests.

KEY WORDS: Concrete, FEA, Nonlinear analysis, Prestressed, Strands, Torsion.

الخلاصة

تقدم هذه الدراسة بحثا عمليا ونظريا لتقصي سلوك اللي للعتبات الخرسانية المسلحة باستخدام حديدتسليح مسبق الاجهاد فقط (ظفائرتسليح عالية الصلادة) وذات مقطع مستطيل الشكل وبدون حديد تسليح إعتيادي. تم إعداد أربع عتبات خرسانية، اثنتان منها مسلحة باستخدام ستة ظفائر مركزية الموقع ومجهده مسبقا لاحداث تأثير قوى محورية مركزة على طول العتبة، والعتبتان الاخريان بدون أي نوع من انواع التسليح وتم فحص كل العتبات المعده في المختبر وباستخدام خرسانة عالية المقاومة تصل الى (٤٤ ميكاباسكال). اشارت النتائج المختبرية للفحص لفروقات زيادة ملحوظة في قيم مقاومة اللي تصل الى ٧٠% في العتبات المسلحة باستخدام الحديد المسبق الاجهاد عنه في العتبات الخالية من التسليح. إن نوع التشققات الناتجة تم تتبعها خلال مراحل الفحص المختبري وملاحظة نوع الفشل الحادث وسلوكية العتبات تحت

تأثير الاحمال المسلطة. تم استخدام برنامج الشعرباف في مقارنة النتائج المحصول عليها مختبريا من الناحية النظرية ووجد تقارب مقبول بين حالات الفحص ونتائج التحليل النظري حيث تم اعتماد نموذج طابوقي مكون من عشرين عقده لغرض تمثيل الخرسانة واعتمد العنصر المطمور لتمثيل عناصر حديد التسليح المستخدم، إضافة لادخال تأثير القوى المحورية في نفس العقد الممثلة لتلك العناصر الحديدية.

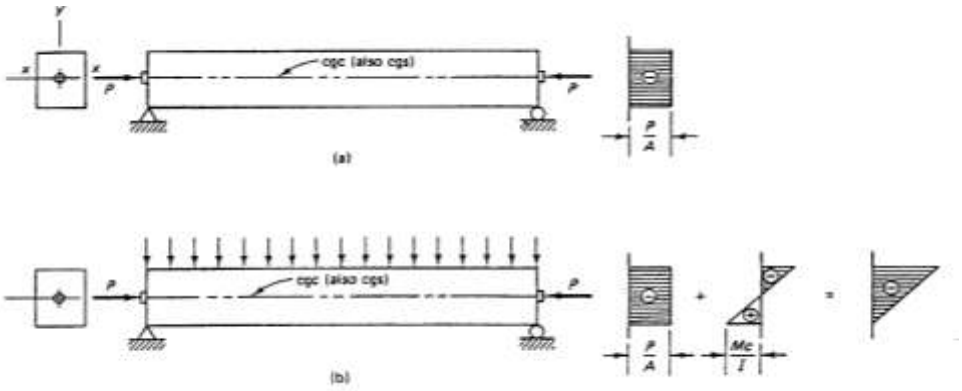
INTRODUCTION

Prestressed concrete beams are used extensively in long span and other structures. The simplest form is the rectangular shape. Prestressed concrete members do not need large maintenance; a longer working life is possible due to better quality control of the concrete. Very large spans such as in segmental bridges or cable-stayed bridges can only be constructed through the use of prestressing, [Nawy, 2000].

The determination of the stress resultants (due to twisting moments or cross sectional distortion) and deformation in straight and curved, simple or continuous girders is complex and requires specified relationships between geometry, loads and deformations.

Concrete is strong in compression, but weak in tension: its tensile strength varies from 8 to 14 percent of its compressive strength [Lin and Burns 1982]. Due to such a low tensile capacity, flexural and torsional cracks develop at early stages of loading. In order to reduce or prevent such cracks from developing, a concentric or eccentric compressive force is imposed in the longitudinal direction of the structural element. This force prevents the cracks from developing by eliminating or considerably reducing the tensile stresses at the critical midspan and support sections at service load, thereby raising the bending, shear and torsional capacities of the sections. The sections are then able to behave elastically, and almost the full capacity of the concrete in compression can be efficiently utilized across the entire depth of the concrete sections when all loads act on the structure. Such an imposed longitudinal force is called a prestressing force; it is a compressive force that prestresses the sections along the span of the structural element prior to the application of the transverse gravity dead and live loads. Fig. (1).

The prestressing force (P) that satisfies the particular conditions of geometry and loading of a given element is determined from the principles of mechanics and from stress-strain relationship of concrete.



(a) Concentric tendon, prestress only. (b) Concentric tendon, self-weight added

Fig. (1): Concrete fiber stress distribution in a rectangular beam with straight tendon [Lin and Burns 1982]

A prestressed concrete beam shows considerably higher torsional stiffness and higher ultimate torque when compared with a plain concrete beam. This increase in torsional strength by the prestressing force may attributed to the higher inter-friction or interlocking force between the slices of the beam (Coulomb friction theory).The present study will show the extent of this effect.

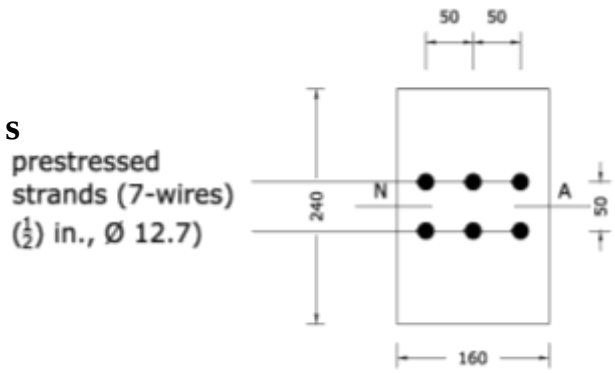
EXPERIMENTAL STUDY

Four beams were tested in this investigation (2 groups): Group one included two samples of concrete beams with six strands only without typical ordinary reinforcement. These two samples were tested under pure torsion only. The strands were distributed on two layers at middle section (two layers) of (3 m) length. Group two include two samples of concrete beams without any type of reinforcement. Both of them were plain concrete of (2m) length and they were tested under pure torsion. Rotations due to applied loads were measured. The strains at many positions on the concrete faces (sides and upper surface) of the beams were recorded. The influence of prestressing force on the increase of ultimate torsion values was investigated. The concrete types (geometry and properties) were kept unchanged for both groups.

TEST SPECIMENS

The nominal dimensions of the tested beams are shown in Fig. (2). A design equation adopted by ACI-318-0٥ Building Code was used to estimate the main factors of concrete strength and the torsional capacities.

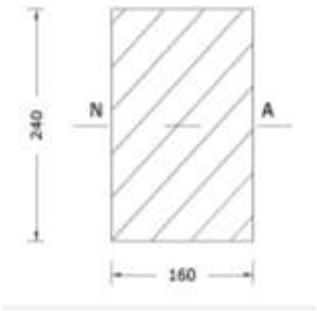
The main reinforcement consisted of (6-strands, 7-wires, ϕ 12.7 mm, 1/2 in.), passing grade 250ksi.



(a)Six strands in cross-section

(b)Isometric view

Fig. (2): Details of test beams (Cross sections and isometric view of the samples, dimensions in mm)



(c)Cross-section of plain concrete beam

Fig. (2): continued

THE BASIC MATERIALS

Prestressed concrete utilizes high quality materials, namely high strength steel and concrete. All the materials were tested under standard specifications. In manufacturing the control and the test samples ,the following materials were used: ordinary Portland cement(Type I);crushed gravel with maximum size of (14mm);natural sand from Al-Ukhaider region with maximum size of (5mm) and fineness modulus of (2.84).The mix proportions for HSC are presented in Table(1).

Prestressing (7-wires) strand has the following components, f_{py} = 1570MPa, f_{pu} = 1770MPa, A_{ps}=93mm², Ø=12.7mm, 1/2 in. and E=195000MPa.

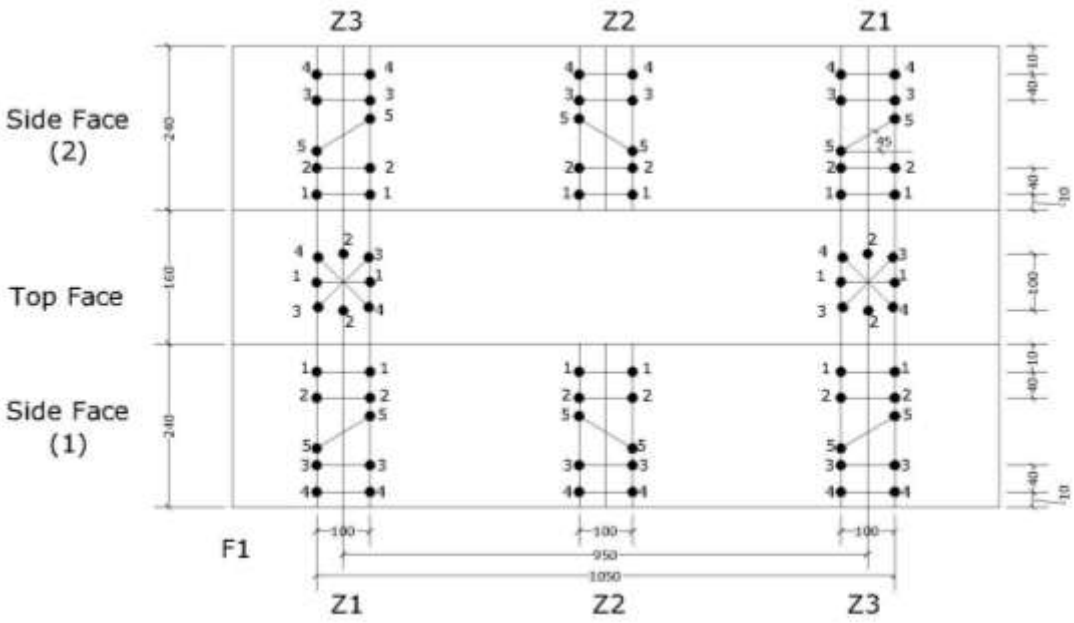
Table (1) Concrete mix proportions

Components	Quantities
Cement type I	500 kg/m^3
Sand (5 mm)	600 kg/m^3
Coarse aggregate (14 mm)	1150 kg/m^3
Water	190 L/m^3
Superplasticizer	0
Water / cement ratio (w/c)	0.38
Strength (cylinder 150*300 mm)	(f'_c) = 35 MPa at 7 days (f'_c) = 42 MPa at 28 days (f'_c) = 50 MPa at 60 days
Density (γ)	2440 kg/m^3

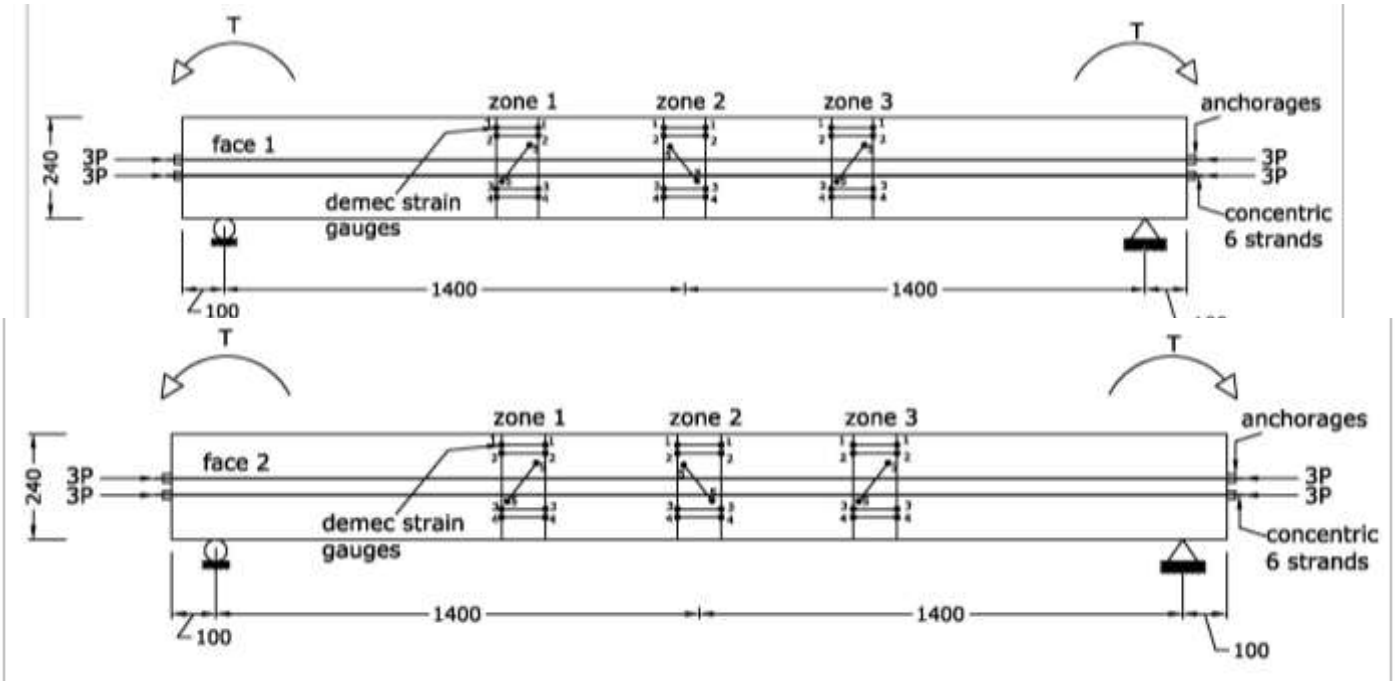
TEST MEASUREMENTS AND INSTRUMENTATION

INCREMENT OF STRAIN MEASUREMENTS (CONCRETE COMPRESSIVE AND TENSILE STRAIN)

The middle zone (1500 mm) of the total length of the specimens was determined as a measuring zone. The longitudinal concrete compressive strain and tensile strain at the extreme top and bottom fibers were measured above and below the neutral axis and at an inclined direction with (45°), using demec strain gages. Demec gage points were placed at a spacing of 100 mm along the length of the measuring zone for all beam specimens, Fig. (3). The demec points can be seen at top face with shape of 4 points as a star, with respect to longitudinal axis or direction of the beam parallel to this axis, and in vertical and inclined directions which represent the expected strain directions under the applied torsional loading conditions, Fig. (3).



The prestressed concrete beams were with full prestressing, testing was under pure torsion, the concrete strain was investigated for compression and tension fibers, Fig. (4).



B. demec strain gage nodes for (side face 2)

Fig. (4): Concentric 6-strand beam under pure torsion with demec strain gages

The strain diagrams for the concentrically prestressed concrete beams, represent the strain variation over the fibers and the behavior of concrete under pure torsion case (full prestressing), these are shown in Figs. (5) to (12).

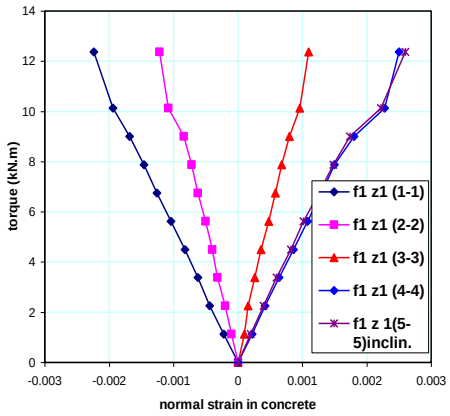


Fig.(5):Torque-strain diagram in concrete (face1/zone1)

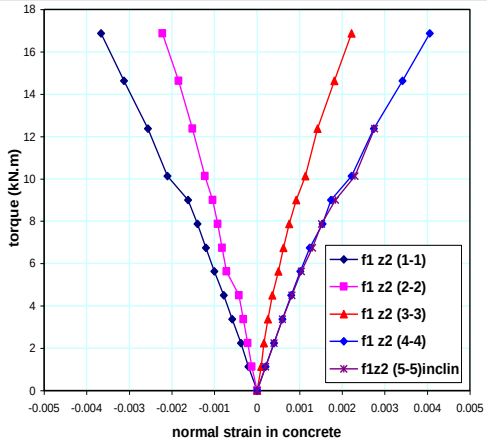


Fig.(6):Torque-strain diagram in concrete (face1/zone2)

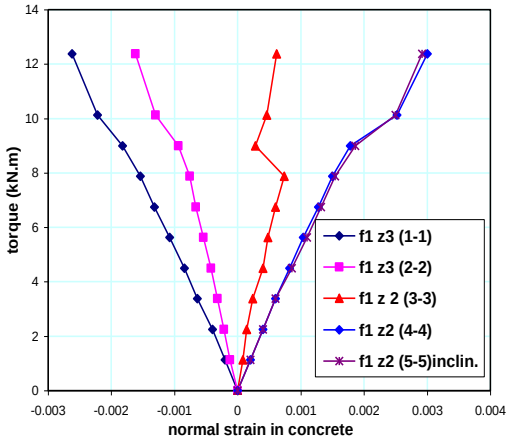


Fig.(7):Torque-strain diagram in concrete (face1/zone 3)

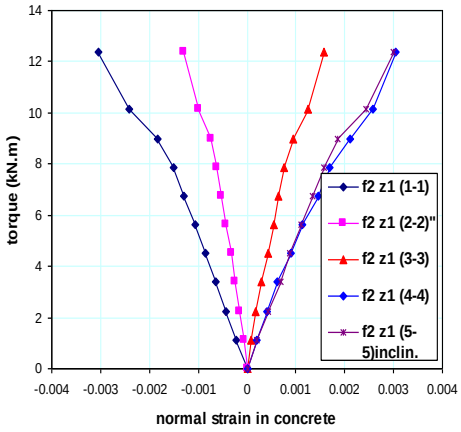


Fig.(8):Torque-strain diagram in concrete (face2/zone1)

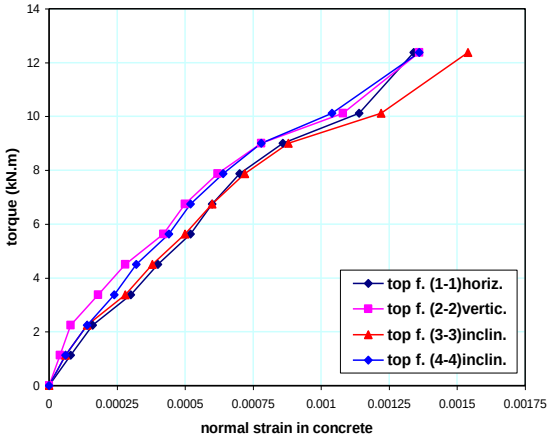


Fig.(12):Torque-strain diagram in concrete (top face,node 2)

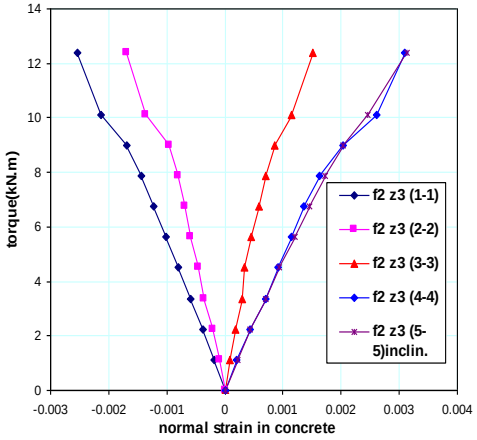


Fig.(10):Torque-strain diagram in concrete (face 2/zone 3)

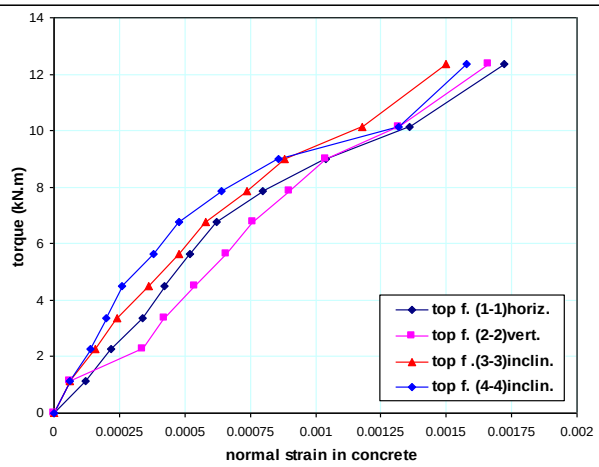


Fig.(11):Torque-strain diagram in concrete (top face,node 1)

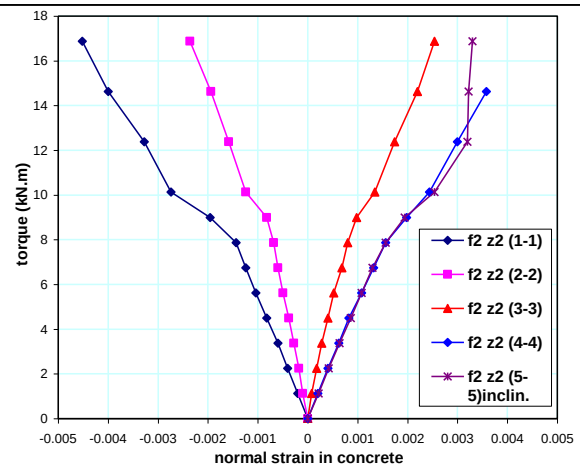


Fig.(9):Torque-strain diagram in concrete (face2/zone2)

ANGLE OF TWIST MEASUREMENTS

For the important application of prestressed concrete beams under the proposed type of this loading condition, pure torsion, the investigation of twisting angle is essential as a relationship with torque value, and also for estimating the torsional capacity of the prestressed concrete beams. A simple mechanical system was arranged with two dial gages in order to record the upward and downward deflection due to the effect of load (torque) on rotating the beam sample cross-section as shown in Fig. (13), within the increasing applied load.

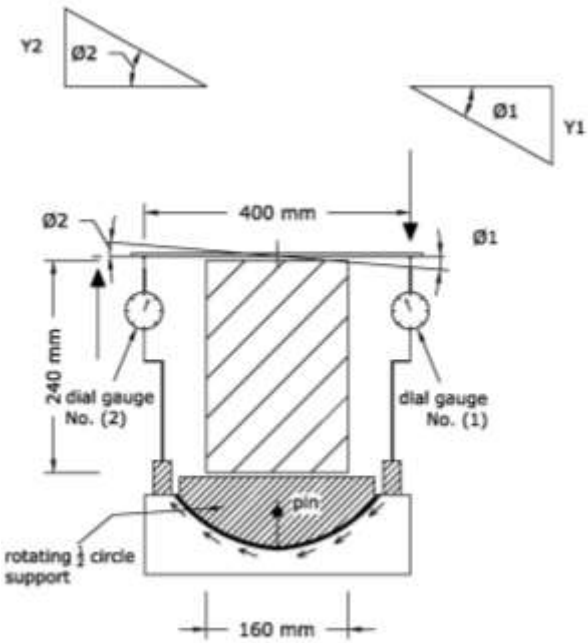


Fig. (13): Angle of twist instrumentation

TEST PROCEDURE OF BEAM SPECIMENS

All beam specimens were tested as simply supported over an effective span of (2800mm) with symmetrical boundary conditions as shown in Fig. 14. Test started with the application of 2kN load to set and check the dial gages, then unloading to zero. At zero loading, initial readings of dial gages and mechanical strain gages (demecs) were obtained. The load was applied in (10 to 25) stages depending on the type of sample and loading conditions. At each loading stage, all the dial gages and strain gage (demec points) readings were taken. The interval between two consecutive stages was roughly 10 minutes. The overall testing time took an average 2 – 2.5 hours, depending on the deformation capacity of the beam tested. The load was continued until failure (defined as the highest capacity beyond which loading dropped) takes place, Fig. 14.

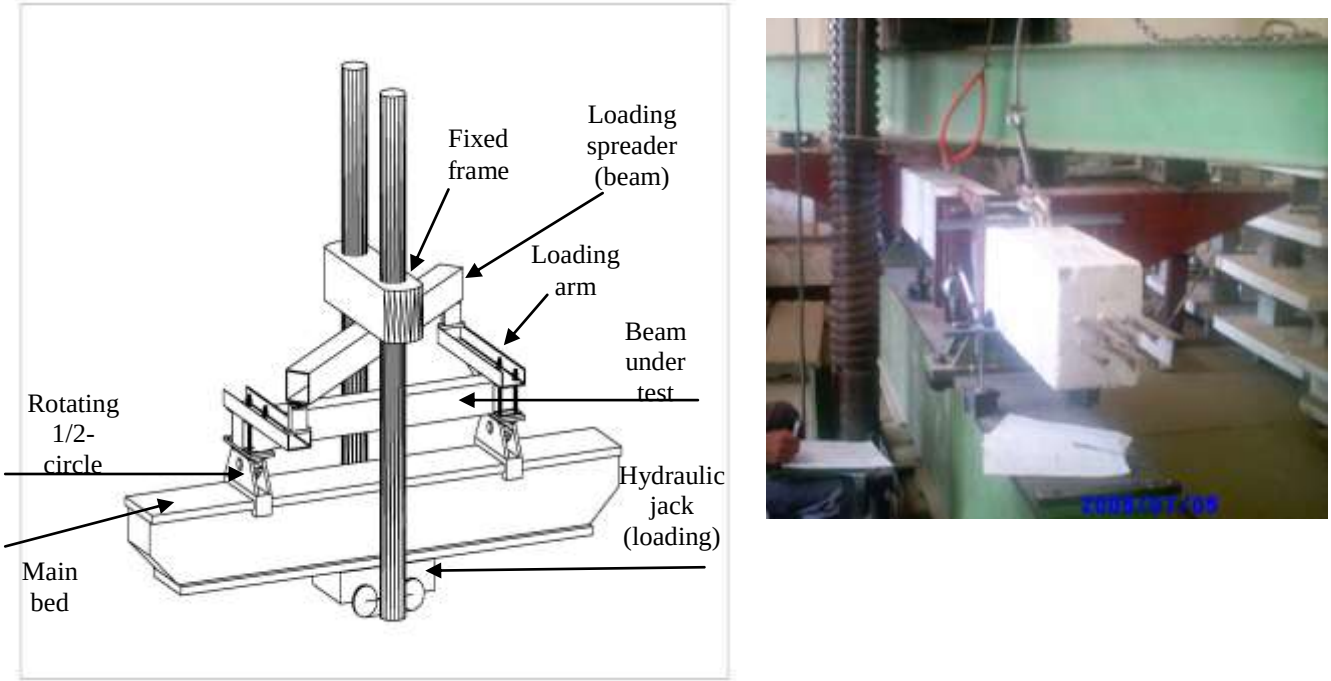


Fig. (14): Concentric prestressed concrete beam sample under testing

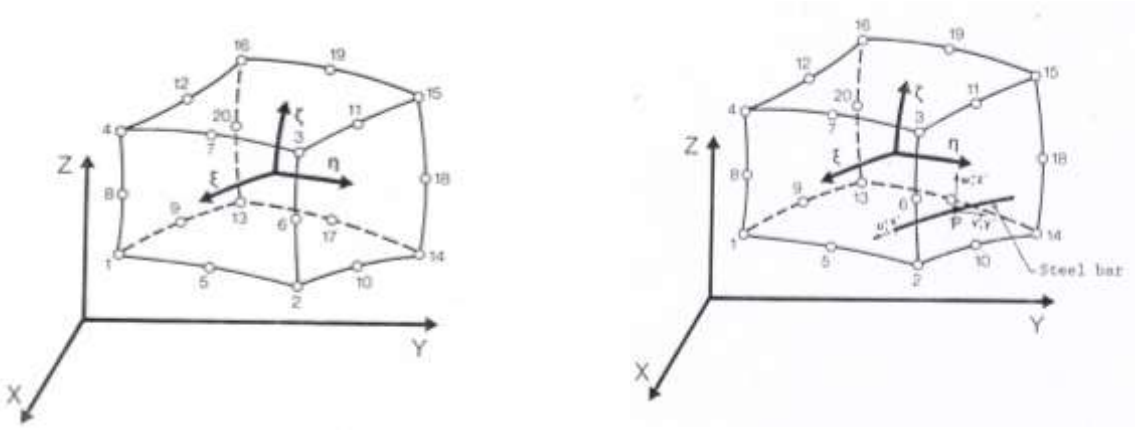
THEORETICAL STUDY

In order to study the behavior of the prestressed concrete rectangular beams under the effect of torsion and to compare the results with the all tested beams, the nonlinear finite element program P3DNFEA [1] was used for this purpose.

FINITE ELEMENT IDEALIZATION

In the present study, the chosen segments were modeled by using the quadratic 20-node brick elements as shown in Fig (15), for all length of the specimen (3m), in order to consider the

effect of initial prestressing forces along the actual length. Longitudinal reinforcement was simulated as embedded one-dimensional elements in the brick elements. Since this study is devoted for the analysis of prestressed concrete beams, so more efficient prestressed models capable of being incorporated in a finite element solution were used.



(a) 20-Node isoparametric brick element [1]

(b): representation of reinforcement (Embedded representation)

Fig. (15): Concrete idealization and reinforcing bar representation

In this approach, the equivalent load method is used to compute the force applied by the tendon upon the structure where the prestressing tendon takes the form of a particular loading case and as such it may be taken into consideration in the form of appropriate local loads at the level of each element. Practically, this means that the effect of prestressing manifests itself in the evaluation of the vectors of primary nodal forces only.

The method is adopted here with the consideration that the tendon has straight profile along the beam and is fully bonded (prestressing by pretensioning) (no inclined or parabolic profile and without losses in tendons) and so the prestressing reinforcement was represented as embedded one-dimensional elements, with local coordinates. The straight tendon being in one level along the beam specimen.

MODELING OF BEAMS

The actual dimensions of the tested beams with special reinforcement and loading conditions are shown in the figures below. Each two beams of group one (G1), had length (3m) and (160x240) mm cross-section, with concentric prestressing reinforcement only, and with no ordinary reinforcement, while the two beams of group two (G2), had length (2m) and

(160x240)mm cross-section, also without any type of reinforcement (plain concrete specimens). Both types of beams were tested under pure torsion to investigate the torsional capacity and the response by the angle of twist parameter and also the magnitude of the applied torque over these samples with effect of prestressing reinforcement, see Table (2).

Table (2) Classification of concentric prestressing beams of plain concrete group with its properties under pure torsion.

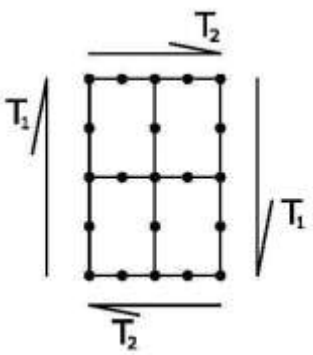
MESHING AND LOADING CONDITIONS

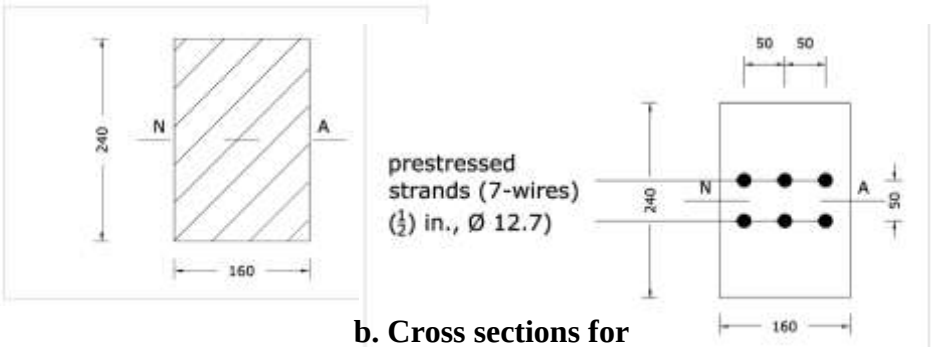
The behavior of the beams under torque was investigated using the angle of twist and the ultimate failure torque.

Group No.	Beam No.	f'_c (MPa)	Prestressing reinforcement No. of stands			Ordin. Reinf.	f_{py} (MPa)	X-section	
			No.	A(mm^2)	dia.			b(mm) width	H (mm) depth
prestressed concrete beams (6 strands)	B1	42	6	558	12.7	None	1570	160	240
	B2	43	6	558	12.7	None	1570	160	240
plain concrete beams	B3	44		None		None	None	160	240
	B4	43		None		None	None	160	240

The geometry and the finite element mesh and with the applied loading are shown in Fig. (16). The whole beam is modeled with 56 brick elements and a total number of 441 nodal points, as shown below.

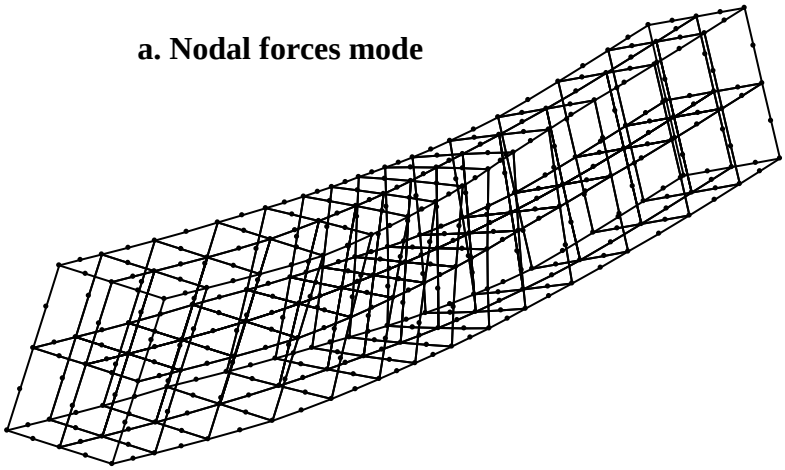
Angle of twist will decrease due to initial prestressing force (as external axial load) if existing.

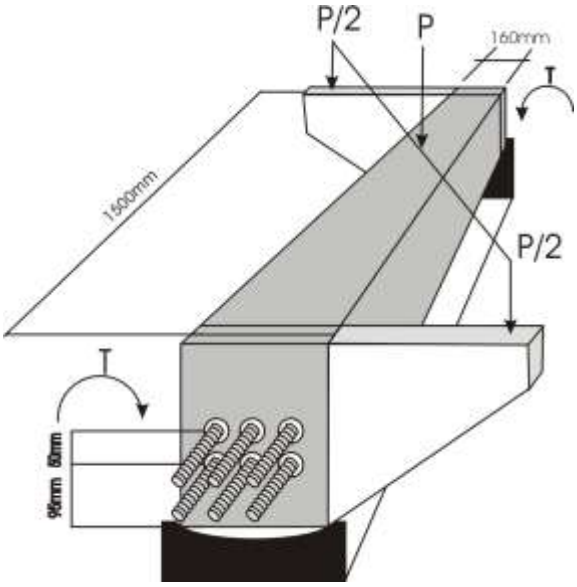




b. Cross sections for prestressed and plain concrete beams

a. Nodal forces mode





c. Finite element mesh

d. Pure torsion case (concentric prestressing and plain concrete)

Fig. (16): Deflection of concrete specimen during test

RESULTS AND DISCUSSION

In Fig. (17), the torque-angle of twist relations of the concentric prestressed concrete beam with 6 strands is shown. While in Fig

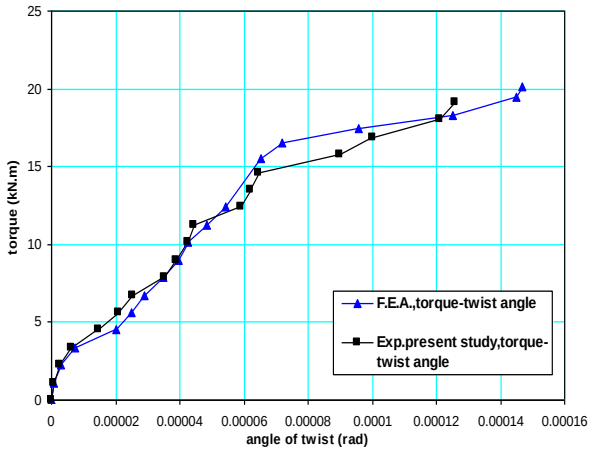


Fig.(17):Experimental and analytical torque -angle of twist diagram (concentric 6-strands prestressed beam under pure torsion)

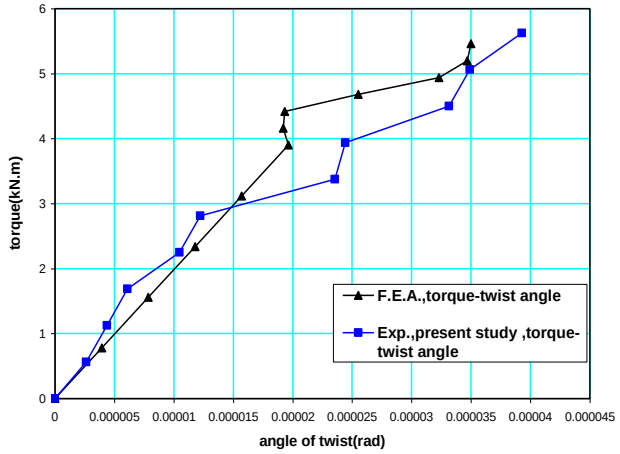


Fig.(18):Experimental and analytical torque-angle of twist diagram (plain concrete beam under pure torsion)

(18), the torque-angle of twist relation of the plain concrete beams is shown. Good agreement is obtained with experimental results. The values of the angle of twist decrease due to existence of initial prestressing force (as external axial load).

GENERAL BEHAVIOR

All beams that were tested under pure torsion loading failed in torsion. At early loading stages, torsional cracks (spiral or helical cracks) were observed first in the middle zone of the beam (pure torsion region). As the load was increased, more torsional cracks developed in the pure torsion span. A main single crack with helical mode was generated at the midspan and gradually extending from side faces and from top and bottom of the sample. Clear cracks at the compressive region of the beam (top face) where the compressive stresses exist was found to increase due as the torsion load increases.This mode of failure associated with a diagonal tension cracks inclined at 45°, as shown in Fig. 19. Sudden failure (sudden seperation) occurred. The cracking patterns are similar for the samples of concentric prestressed members and the samples of plain concrete, with inclined angle approaching 45°. However the main difference is the drop of the failure load (torsional resistance) between prestressed beams and the beams of plain concrete, where presence of prestressing strands in the section reduces the angle of twist of concrete sample and thus increased the ultimate torsional resistance.

Fig.(19):Failure of plain concrete beam
under pure torsion loading

FAILURE MODE

The failure modes for the tested beams were the formation of diagonal tension crack (due to torsion load), as shown in Fig. 20. In table (3) the differences between experimental and theoritical results are presented.

(a)Single diagonal crack mode failure for plain concrete sample

(b)Single helical crack for prestressed concrete beam with

Fig.(20): Modes of failure for plain and Prestressed concrete beams under pure torsion loads

Table (7.8): comparison of experimental and analytical results for all samples

Group No.	Beam No.	Type of load	Tult. Exp. (kN.m)	Tult. (max) F.E.A. (kN.m)	ult. Twist angle (θ) rad. Exp.	ult. Twist angle (θ) rad. F.E.A	% of diff. Tult. (kN.m)	% of diff. in (θ) rad.
Group 1 (6-strands)	B1	(T)only	**19.125***	20.125	$1.26x10^{-4}$	$1.467x10^{-4}$	3	14
	B2	(T)only	**19.125**	20.125	$1.26x10^{-4}$	$1.467x10^{-4}$	3	14
Group 2 (plain concrete)	B3	(T)only	**5.625***	5.46	$3.926x10^{-5}$	$3.207x10^{-5}$	1	18
	B4	(T)only	**5.625**	5.46	$3.926x10^{-5}$	$3.207x10^{-5}$	1	18

***The main difference in values of ultimate torque that would be applied over the beams for plain and concentric prestress reaches to 73%.**

CONCLUSIONS

Based on the results obtained from the experimental work and the finite element analysis for concentric prestressed and plain concrete beams, the following conclusions are presented:

1. Presence of the prestressed strands delays concrete failure and leads to higher failure load.

2. The ultimate torsional resistance is increased by 73% for the tested beams with prestressed strands relative to plain concrete beams.

3. The controlling crack propagation, the rate of cracks widening and the load carrying capacity are higher for the prestressed concrete beams relative to plain concrete beams.

4. The nonlinear finite element method utilized in this study has shown to be capable of reproducing the experimental response of the prestressed concrete beams. The isoparametric brick elements with embedded steel bars proved to be suitable for predicting the state of ultimate load, deflections and stress with good accuracy. The differences with experimental values (in deflection or ultimate load) were in the range (1-18%).

5. The effect of providing axial compressive force by prestressing was significant to get higher torsional resistance than in plain concrete beams. This phenomenon appeared pronounced in concentric prestressed concrete beams, where the ultimate torque increased considerably (more than 3 times) due to the applied compressive stresses on the sections.

NOTATIONS

T_u	**1. Ultimate torsional moment**
f'_c	**Compressive strength of concrete**
T_{uo}	**Ultimate torsional capacity of beam subjected to pure torsion**
A_{ps}	**Area of prestressing steel**
A_s	**Area of reinforcing steel**
E_c	**Modulus of elasticity of concrete**
E_{ps}	**Modulus of elasticity of prestressing steel**
f_{py}	**Yield strength of prestressing bar**
x,y,z	**Global coordinate system**
x', y', z'	**Local coordinate system**
ε	**Strain**
σ	**Stress**

τ	**Shear stress**
ξ,η,ζ	**Curvilinear coordinate system**

REFERENCES

- - Al-Shaarbaf I., **"A Nonlinear Three-Dimensional Finite Element Analysis Program for Steel and Reinforced Concrete Beam in Torsion"**, Ph.D. Thesis, University of Bradford, U.K., 1990.
- - Lin T.Y. and Burns N.A., **"Design of Prestressed Concrete Structures"**, Third Edition, John Wiley and Sons, 1982.
- - Nawy E.G.,"**Prestressed Concrete**", A Fundamental Approach, 3rd Edition, Florida, 2000.

SETTLEMENT REDUCTION UNDERNEATH SURFACE CIRCULAR FOOTING RESTING ON REINFORCED SOILS.

Hayder M. Mekkiyah
The University of Baghdad, Civil Engineering Department.

ABSTRACT

An analytical approach, adopted to find the settlement of foundations resting on reinforced soil based on the test results on a model surface circular footing resting on reinforced soil, is summarized. The soil was reinforced using biaxial geomesh. The settlement was determined by considering the compatibility of strain (settlement) between soil and reinforcement element underneath the foundation. Theoretical equations were used to estimate the settlement either from the superstructure loads or from in-situ plate load tests on the reinforced soil system. The type of geomesh used in this study has been determined based on the grain size distribution of the soil. The investigation in this study used two different types of geomesh. Uniformly graded sand was used to make it easier to control the density and fabric in different tests. It was found that initial horizontal and vertical movement of the reinforcement is needed to mobilize the reinforcing strength. Further, the initial settlement at small loads could be avoided when the reinforcement was placed closer to the base of the footing and there was an improvement in the bearing capacity value of the footing. When the reinforcement is placed away from the base of the footing(greater than B), the initial settlement decreased with a slight improvement in the bearing capacity compared with that of unreinforced soil. Non-dimensional factors were developed for settlement calculations based on the experimental test results from a series of laboratory tests on the model footing. Additional tests were performed on the model footing to evaluate the effect of the number of reinforcement layers and the depth of the top most reinforcement layer on the settlement and the improvement in the bearing capacity of the footing-reinforced soil system.

الخلاصة

اقترح في هذه الدراسة تحليل لحساب الهبوط تحت الأسس المسلحة ترابيا مبنيا على تجارب الفحوص المختبرية لهذه الأسس الجالسة على التربة المسلحة حيث أن التربة تم تسليحها بواسطة قطع (شبكة) تسليح بلاستيكية. كان مبدأ التحليل مستندا إلى التوافق في الانفعال الحاصل تحت الأسس لكل من التربة وقطع التسليح.أن المعادلات المقترحة لحساب الهبوط تحت هذا النوع من الأسس يمكن استخدامها لكل من نتائج الفحوصات الحقلية أو أية أحمال متوقعة من المنشأ يراد حساب الهبوط تحتها ،تم استخدام رمل منتظم متدرج في هذه الفحوص .لوحظ من

أجل حصول حركة (رأسية / أو أفقية) تحت الأسس المسلحة فلابد من وجود أحمال أكبر من تلك المقارنة بالأسس غير المسلحة ، وأكثر من ذلك أن هذا الهبوط الأولي تحت هذه الأسس يمكن إهماله عندما قيم الأحمال الأولية المسلطة قليلة وشرائح قطع التسليح قريبة من قاعدة الأسس كما لوحظ أيضا تحسن كبيرفي تحمل الأسس لهذه الحالة من التحميل ، في حين أن هناك تحسن أقل سعة التحمل عندما تكون هذه الشرائح بعيدة عن الأسس وفي المقابل الهبوط المتوقع يكون أكبر ،أن مثل هذا التباين في مقدار الهبوط لمثل هذه الحالات من التحميل تم أخذه بعين الاعتبار حينما تم وضع المعاملات اللابعدية من أجل حساب الهبوط المتوقع تحت الأسس في المعادلات المقترحة في هذه الدراسة.تم حساب الهبوط لحالات فحص أخرى غير تلك التي أجريت في البحث للتأكد من مقدار الهبوط المتوقع فكان مقدار هذا الهبوط يتناسب وطبيعة التربة ومقدار الأحمال المسلطة على هذه الحالة.

KEY WORDS:-Settlement, Circular, Footing, reinforced, Soil, Reduction

INTRODUCTION

Reinforced earth technique is one of the most promising materials that have emerged in the last 30 years from intensive research that has been carried out into alternative construction materials. Reinforced earth technique is not new, the earliest remaining examples of soil reinforcement are ziggurat of ancient city of Dur-Krigatzu in Iraq (6000 B.C.), and the great wall of China. It is also known that Romans have used earth reinforcement technique (Ignold 1982). Further, there are a limited number of studies in the literature on the possibility of using analytical developed equations to estimate footing settlement resting on reinforced sand. This paper reports the initial findings of such a study and attempts to provide a relatively simple approach to estimate the settlement of circular footings resting on reinforced sand. The proposed approach is based on test results of a model circular footing.

PHYSICAL MODELING

Loading tests of a model circular footing resting on the surface of a reinforced sand subgrade were performed using steel lever-arm system. The sand is placed in a square wooden box of internal dimensions of 570 mm x 570 mm x 800 mm, and 10 mm in thickness, stiffened by means of steel strips and the inside of the box was covered with two sheets of polyethylene. The model footing consists of circular aluminum metal with a diameter of 50 mm and a thickness of 50 mm. The sand flows in a flexible hose through sieve No. 4 and then to the box, where the falling height of sand was fixed at 600 mm. It was found that pouring the sand from this height in 25-mm lifts, gives a unit weight of 18.8 kN/m^3 and a relative density of 65% (medium dense sand) this lifts was kept constant for the whole layers in all tests (similar to that recommended by Bieganousky et. ,al 1976 was used).The testing were carried out at Baghdad University.

MATERIALS USED

The sand used are passing sieve No.4 was washed with running water to remove dust as much dust as possible. The sand was then air dried before before sored in barrel,sieve analysis was carried out and a grain size distribution curves was obtained The uniformity coefficient of sand was determined as 3.2 .Laboratory tests were carried out on the sand to get some other properties and these values are listed below:

Specific gravity ……Gs=2.63

Maximum unit weight (γd max)= 19.71 KN/m^3
Minimum unit weight (γd min)= 16.46KN/m^3
Unit weight in the Box = (γd)= 18.8 KN/m^3
Void ratios Calculated based on (γd max and γd min):
emax = 0.567 emin = 0.309

The angle of internal friction was determined as 39° using Triaxial tests. The reinforcement used in the tests were of polymer geomeshes commercially known as Netlon geomesh (CE111 and CE121) having an aperture size of (8 mm x 6 mm) and thickness of 2.90 mm and 3.3 mm for CE111 and CE121 respectively with dimensions of 540x540 mm. The tensile strength was 2.0 and 7.68 kN/m respectively. The bearing capacity and settlement of the footing resting on sand depend on properties of sand such as the angle of internal friction ϕ and the relative density, size, shape and embedment depth of footing (Lambe and Whitman, 1979). The results obtained from small scale model tests such as the one used in this study are usually hindered by limitations associated with size and boundary effects. As a result, it is of importance to keep such limitations in mind when designing such small scale model tests and when interpreting and extrapolating results to full scale footings.

TEST RESULTS AND DISCUSSION

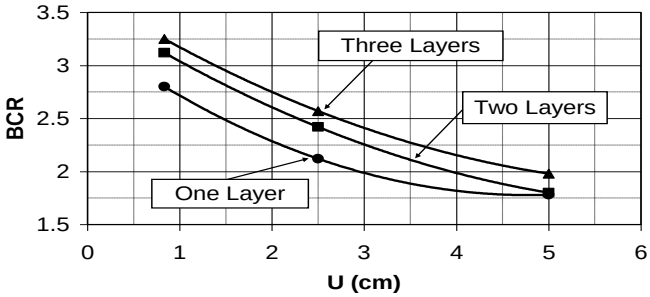
The following parameters were considered in this study:

a. Depth of top most reinforcement layer.
b. Number of reinforcement layers.
c. Improvement in the subgrade reaction value.

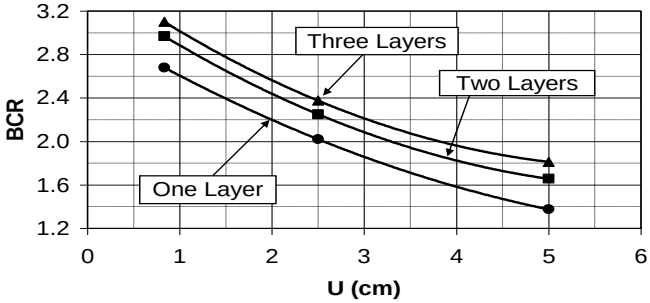
The bearing capacity of the footing-soil system with and without reinforcement, q and q_o respectively, were obtained from load-settlement relationship. The bearing capacity ratio (BCR), which is defined as (q/q_o) and represented at a settlement of 5% of footing width , most of failure cases (load -settlement curves) capacities starts within this value .

A-EFFECT OF TOP MOST REINFORCEMENT LAYER

Tests were carried out to investigate the effect of distance between the footing and the top most reinforcement layer (U), where U is used as a ratio of the diameter of the footing D. The relationship between U and bearing capacity ratio BCR is drawn for different values of U (i.e U=D/6, D/2, D &ΔHused=2cm between reinforcement layers), different types of reinforcement (Netlon CE111, and CE 121), and for different number of layers (N) as shown in Figures (1 and 2), as example of one layer layout below footing different tests were carried out at different (U) locations and similar for two and three layers of reinforcements. It is found that BCR increases as U decreases for all number of reinforcement layers. As expected and observed, the results show that the soil deformation occurs first in the upper layer of sand, just below the footing, and then propagates to deeper areas as the load is increased. This is due to the existence of the high stress zone below the footing which reflects the benefit obtained by placing the reinforcement at this zone. This observation is in agreement with that reported by Akinmusuru and Akinbolade (1981).



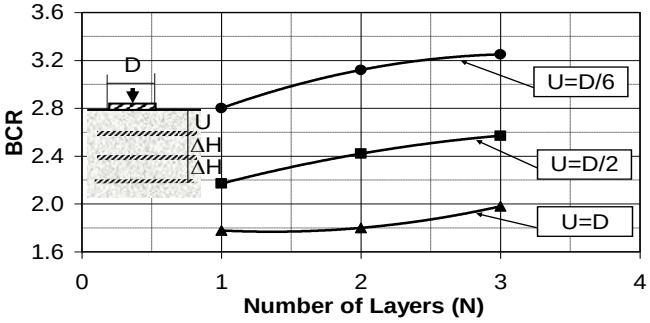
Fig(1) BCR-U Relationship for Netlon CE121.



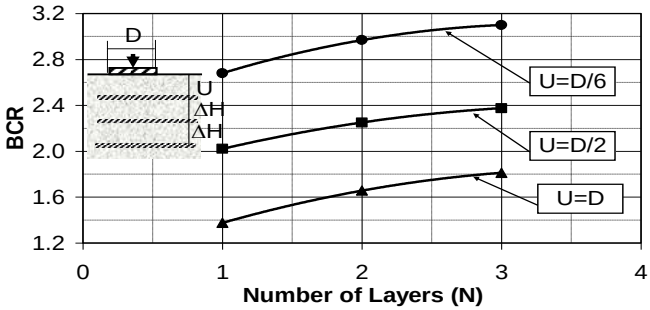
Fig(2) BCR-U Relationship for Netlon CE111.

B-EFFECT OF THE NUMBER OF REINFORCEMENT LAYERS

The effect of increasing the number of reinforcement layers (N) on the bearing capacity of footing is shown in Figures(3 and 4). The gain in the bearing capacity with the number of layers is expressed in terms of bearing capacity ratio (BCR). Figures (3 & 4) show that BCR increases with the number of reinforcement layers for both types of reinforcement(ΔH=2cm). A similar conclusion was reported by Akinmusuru and Akinbolade (1981).



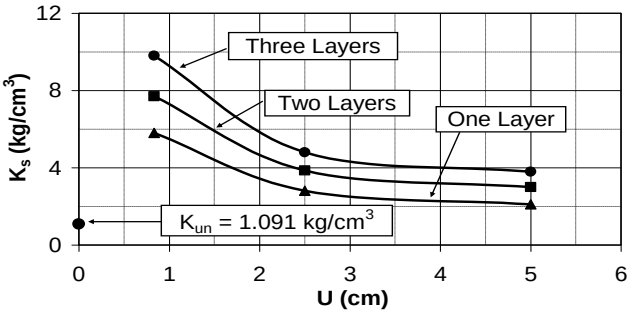
Fig(3) BCR-Number of Layer Relationship for Netlon CE121.



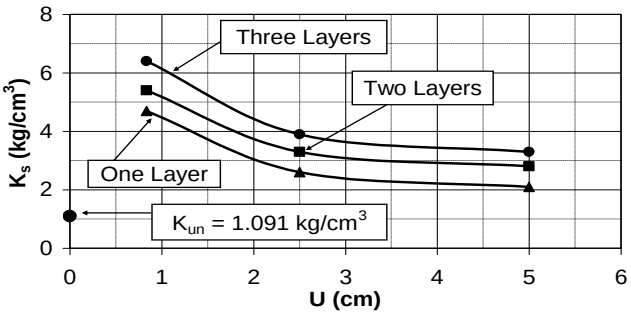
Fig(4) BCR-Number of Layer Relationship for Netlon CE111.

IMPROVEMENT IN THE SUBGRADE REACTION VALUE

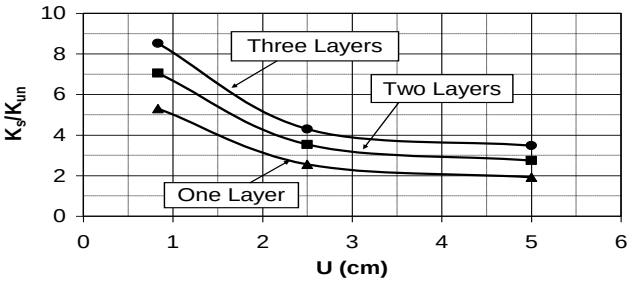
It is concluded from the tests that reinforcing the soil causes improvement in the soil subgrade reaction value (Figures 5-8) these values are calculated from ($\Delta q/\Delta H$) at a settlement of 5% of footing width. . The figures show that the modulus of subgrade increases as U decreases and as the number of layers increases. It is recommended to place the first layer of reinforcement in the zone of initial strain (i.e., close to the footing base at a depth that is less than or equal to B/6), and the second layer in the lower zone of maximum strain at a depth of 0.4B below the base of the footing. Placing the reinforcement in these levels will significantly improve the subgrade reaction value and reduced the footing settlement. The figures also show that the subgrade reaction reaches a steady value when U is larger than 80% of the diameter of the footing (D). Similar observations were also presented by Al-Dobaissi (1990). The values of subgrade reaction of the unreinforced soil (K_{un}) are also shown in the Figs. 5 and 6.



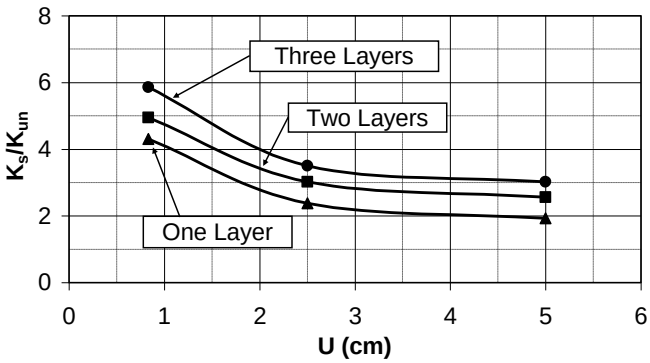
Fig(5) K_s-U Relationship for Netlon CE121.



Fig(6) K_s-U Relationship for Netlon CE111.



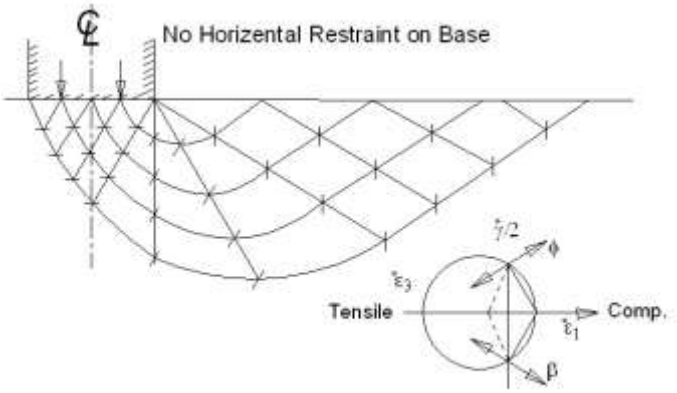
Fig(7)Stiffening Subgrade Factor (K_s/K_{un}) versus U Relationship for Netlon



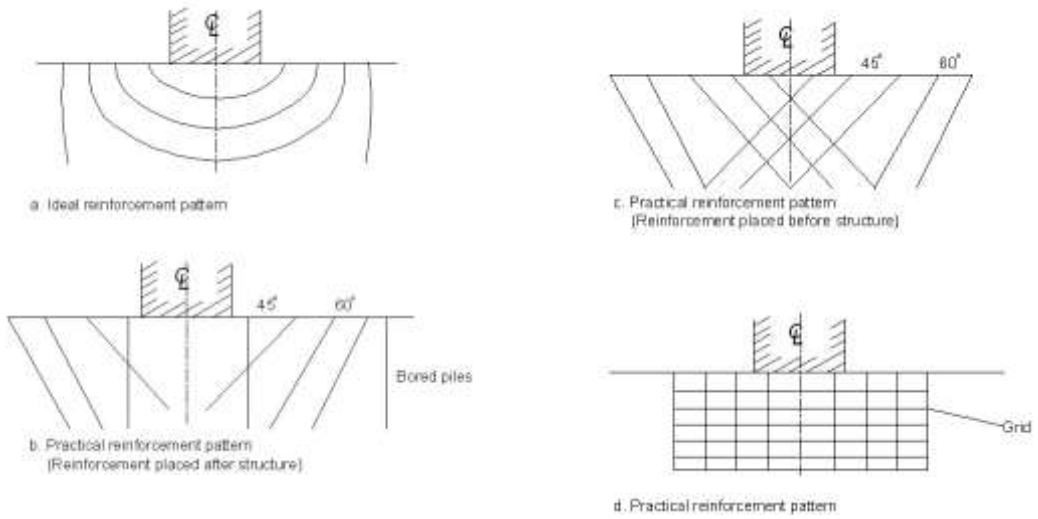
Fig(8) Stiffening Subgrade Factor (K_s/K_{un}) versus U Relationship for Netlon CE111

ALYTICAL AND EXPERIMENTAL APPROACH

The use of reinforcement (Geosynthetics) to improve the bearing capacity of footings and to reduce settlement has been proven to be cost-effective for foundation system. A reinforced soil-foundation system consists of one or more layers of geosynthetics and a control soil placed below a conventional spread footing. The reinforcement is usually placed horizontally. However, there are cases in which vertical or sloped reinforcement may be used below the footing. Further, the reinforcement placed within the tensile arc of strain field causes realignment of the strain field which improves the performance for the load carrying capacity (Colin JFP Jones 1985). The ideal reinforcement pattern for the direction of the principal tensile strain is shown in Figures 9 and 10. As shown in these figures, the ideal pattern has a reinforcement placed horizontally below the footing and becomes progressively more vertical further from the footing (Bassat and Last 1978) .



Fig(9) Zero extension characteristics for dilating soil (After Bassat and Last,1978).



Fig(10) Different reinforcement orientations below the footing (After Colin JFP Jones 1985) .

The calculation of footing immediate settlement for different soil types is estimated on the basis of elasticity, provided that the elastic properties of the soil (modulus of elasticity E, and Poisson's ratio υ) are known. These two parameters can be evaluated in the lab from soil samples obtained during site investigation processes for cohesive soils. However, for granular soils, it is much more difficult, if not impossible in most cases. The in-situ testing on granular soils may not accurately give these soil properties which are needed for the calculation of settlement. In the case of reinforced soil systems, it seems to be difficult to use traditional investigation methods such as borings, or to use other traditional techniques such as pressuremeter tests or cone penetrometer tests. Such methods and techniques require drilling to various depths which will deform the reinforcement mesh below the footing. Plate bearing test on reinforced foundation systems resting on homogeneous sand to a sufficient depth, on the other hand, can be used as an economical alternative. The model footing can be used to estimate the overall modulus of the soil which provides a representative parameter for use in conventional settlement estimation.

From Figures. 5 and 6, it was concluded that the improvement in the modulus of subgrade reaction as a result of reinforcement is in the range of 2 to 10 times that of unreinforced soils. It was assumed that the modulus of elasticity of reinforced soil (E_R) will be increased by the same ratio (i.e., E_R=2-10E_S), where E_S is modulus of elasticity for unreinforced soil and E_R can be estimated from equation (1)

$$E_R = (FI) * Ksun * B (1- \upsilon^2) \text{ ----------(1)}$$

Where:

E_R: Modulus of elasticity for reinforced soil.

FI: Improvement factor (FI = 2 and 10 for 1 and 3 reinforcement layers respectively)

Ks_{UN}: The subgrade reaction value of unreinforced soil.

B: Footing width (for an equivalent square).

υ: Poisson's ratio (recommended ranges are between 0.28 and 0.34 for 3 and 1 reinforcement layers respectively).

While the settlement below a reinforced soil system can be estimated from equation (2) which should be used with the following limitations in mind

- Best estimation for base contact pressure (q) should be used.
- For the circular footing it is better to convert the footing width to equivalent square.
- The sands layer depth can cause settlement to a depth of Z= 1.5 to 2 times B or to a depth where a hard stratum is encountered below the base.

$$\delta_{FlP} = 0.8\frac{q \times B}{E_R} \quad \text{---------- (2)}$$

Where:

δ_{FlP} : Footing and/or plate settlement.

q: Load on footing and/or plate.

B: Footing width (an equivalent square).

When the previous limitations are considered, the settlement estimated from the above equation gives good correlation with the test results.

Another method of analysis was proposed for settlement estimation by adopting a non-dimensional factor for any size of footing or plate size. The value of α factor that will provide a settlement of 25 mm is used in equation (3).

$$\delta_F = \frac{2\delta_P}{\left(B_P / B_F\right)^{\alpha}} \quad \text{---------- (3)}$$

Where:

δ_F : Footing settlement (mm).

δ_P : settlement from a plate test of model footing and/or plate bearing test.

α: **non dimensional factor as shown and proposed in Figures** (14-21).

B_P: plate size (m).

B_F: footing size (m).

By using the plate load-settlement curve for δ_F of 25mm, the value of the corresponding bearing pressure can be found from the curve of the computed value of δ_P from equation (3). This bearing pressure is the safe pressure for a given permissible settlement (δ_F), or one can run a reverse calculation to find out the safe pressure for the settlement criterion. If the footing is allowed to settle for (50 mm) then the value of (α) obtained from Figs. 14-21 should be increased by 20-25%.

YIELD CRITERION IN REINFORCED FOUNDATION SYSTEMS

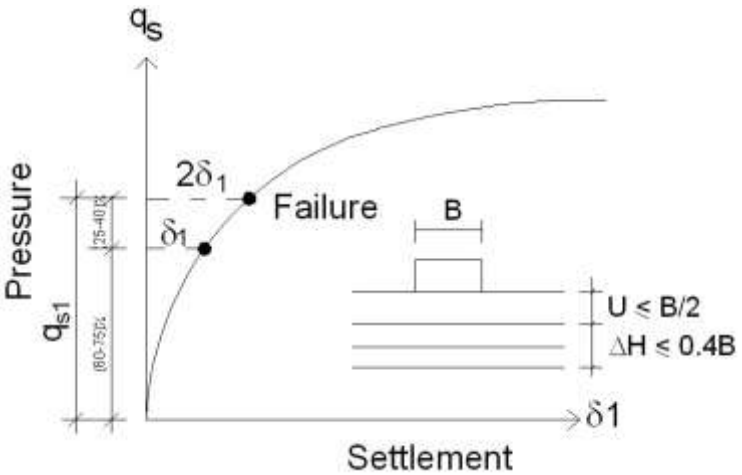
The yield stress is defined when permanent deformation initiates. The yield stress which is a boundary to separate the elastic and plastic deformation for soils is usually not clearly defined and is not a constant value. The locus of the stress at which a soil yields is called yield surface. The stresses smaller than yield stresses cause the soil to respond elastically, and stresses larger than yield stresses cause the soil to respond in an elastoplatic way. The yield stress for soil continuously increases or decreases as the soil hardens or softens. The load settlement curves for reinforced foundation systems was found to be elastic when the reinforcement is placed close to the base of the

footing (i.e., U≤B/2). A number of tests were performed to verify this behavior, and higher yield stresses were obtained at failure due to reinforcement location in this zone (when U is smaller than or equal to B) and due to the inclusion of additional confining stresses in the soil. The additional confining stresses are the result of the placement of the reinforcement in the soil. The proposed failure criterion in the medium dense reinforced sand have been proposed and defined as the bearing capacity at which the settlement is twice the settlement at 60-75% of the safe bearing pressure for the case of U≤B/2 (Fig. 19), Further, the proposed failure criterion in the medium dense reinforced sand has been defined as the bearing capacity at which the settlement is twice the settlement at 80-90% of the safe bearing pressure for the case of U≥ B (Fig. 20).

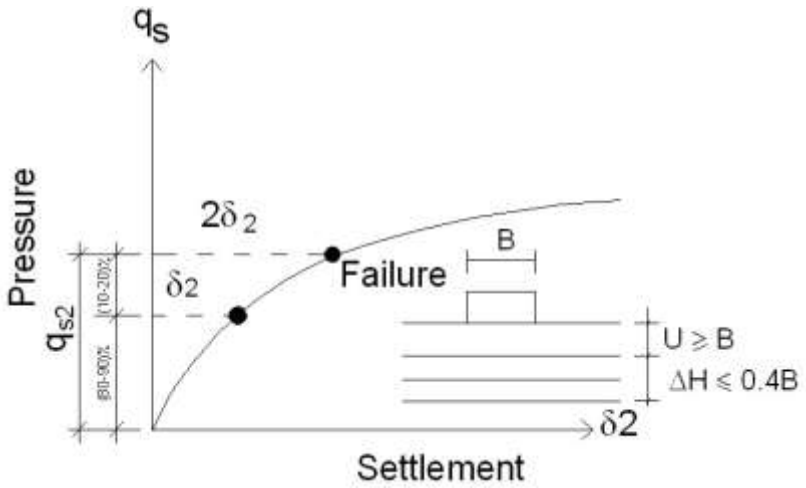
From the tests results it was found that δ_1 is clearly smaller than δ_2, which clearly shows the benefit of reinforcement inclusion in the zone of tension arc, where the zone of high tensile stresses exists. Figure 21 shows the general load settlement trends for both cases.

Additionally, the footing on a reinforced foundation system is more likely to experience a gradual failure curve than a plunging failure. This clearly shows that the settlement is highly reduced when reinforcement is placed closer to the base of footing, while it is improved in a lesser degree when reinforcement is placed further from the footing (Figures 19 and 20).

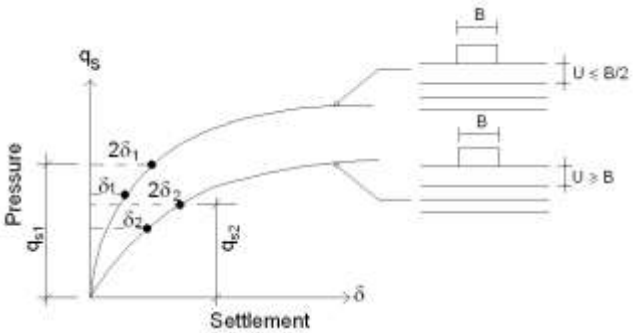
The value of δ_p obtained from equation (3) represents the value $2\delta_1$ and /or $2\delta_2$ in figures 11 and 12 in order to verify the proposed safe bearing pressure in the proposed yield failure criterion for reinforced footing systems. The plate load tests should not be used to determine the ultimate bearing pressure of footings resting on sandy soils because scale effects in such a case give misleading results .



Fig(11) proposed Safe bearing capacity (q_s) for the settlement criterion of circular footing resting on reinforced subgrads (U≤B/2).



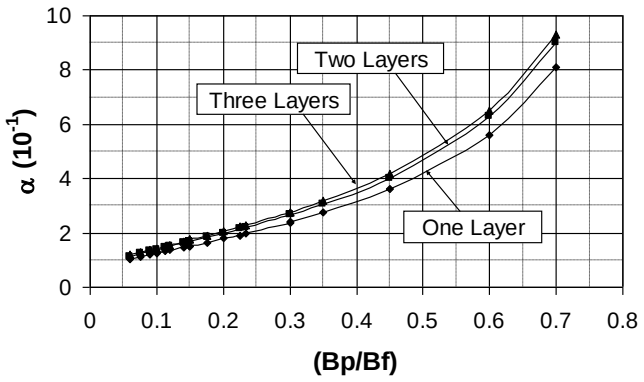
Fig(12) Proposed Safe pressure (qs) for the settlement criterion of circular footing resting on reinforced subgrads (U≥ B).



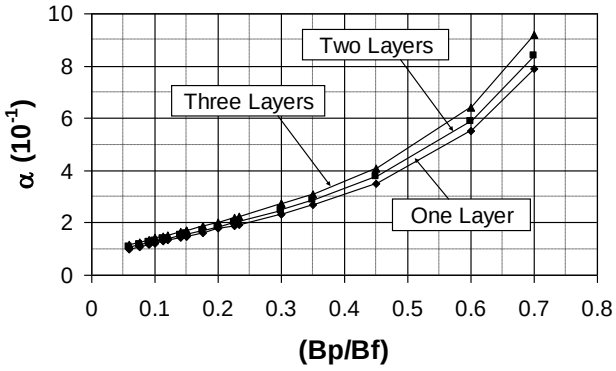
Fig(13) Proposed Safe bearing capacity (q_s) for the settlement criterion for RFS(U≤B/2 and U≥B) rest on sandy soils.

It was also noted that, when the reinforcement was placed in the zone of maximum soil shear, it acted to significantly inhibit the development of a classical bearing failure .

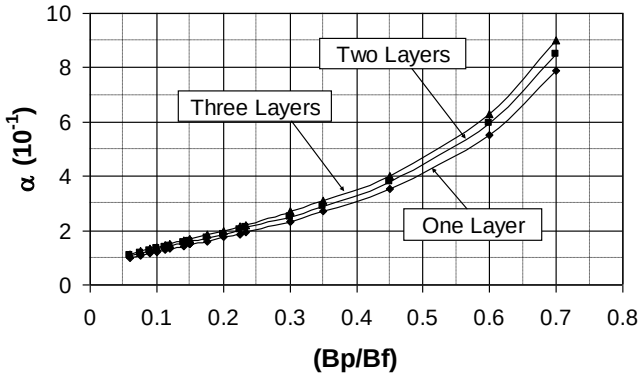
The results in the next figures clearly demonstrate that reinforcement below the shallow footing on sand can reduce the amount of the of settlement, especially differential settlement under the four corners of footings. Footings resting on unreinforced sandy soil settled unevenly, while footings on reinforced soil settled evenly with no tipping of any corners during the observation for the settlement values at the corners after ending the test.



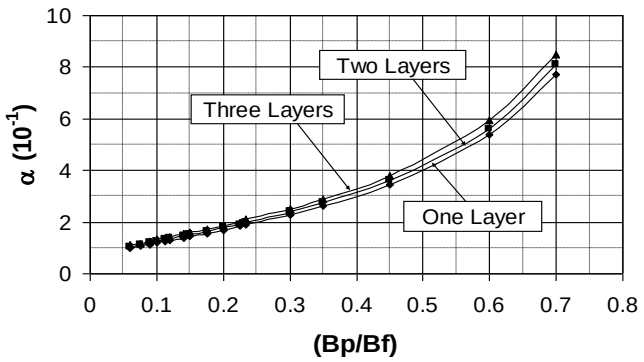
Fig(14) α - (Bp/Bf) relationships for (U=B/6) CE111.



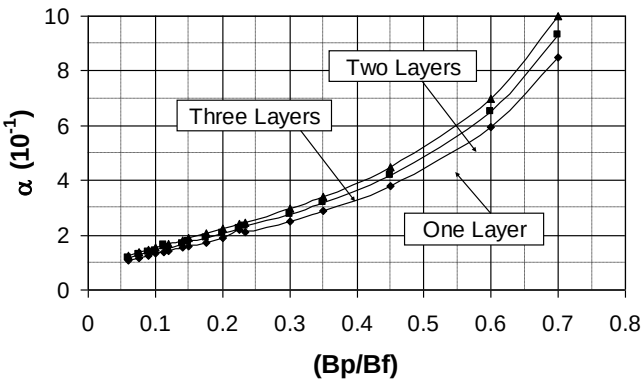
Fig(15) α - (Bp/Bf) relationships for (U=B/3) CE111.



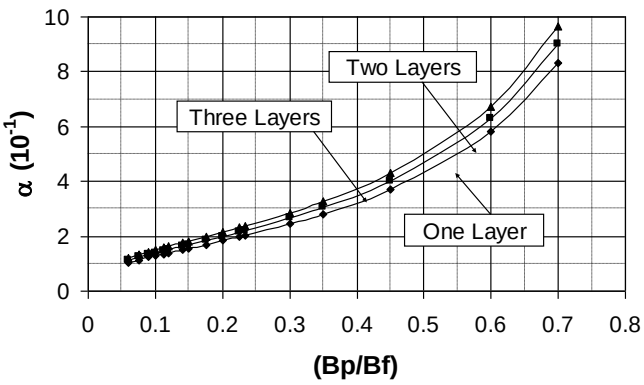
Fig(16) α - (Bp/Bf) relationships for (U=B/2) CE111.



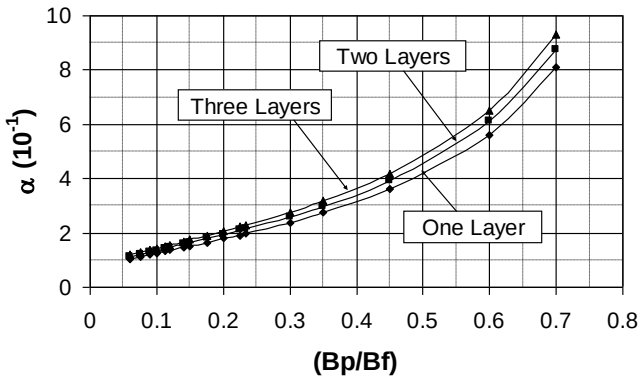
Fig(17) α - (Bp/Bf) relationships for (U=B) CE111.



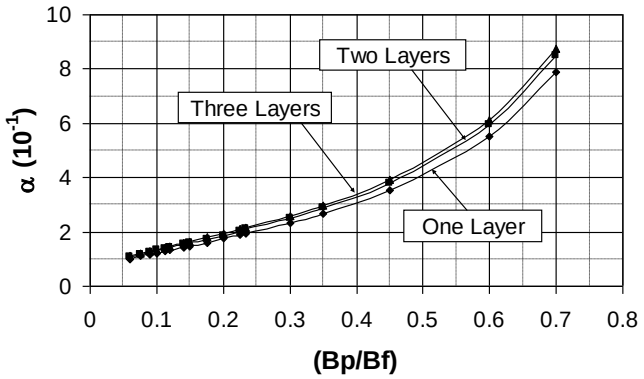
Fig(18) α - (Bp/Bf) relationships for (U=B/6) CE121.



Fig(19) α - (Bp/Bf) relationships for (U=B/3) CE121.



Fig(20) α - (B_p/B_f) relationships for (U=B/2) CE121.



Fig(21) α - (Bp/Bf) relationships for (U=B) CE121.

CONCLUSIONS

The following main conclusions are drawn from this study:

- The depth of top most reinforcement layer is found to be more effective when it is located near the base of the footing (tension arc zone) and the BCR increased up to 3 rapidly when the value of U is close to footing base and the number of layers of reinforcement is three, while a little improvement was achieved beyond that number of layers.
- The settlement value is smaller when a stiff geogrid (CE121) is used below the footing compared with another geogrid (CE111). This was the result of the fact that the value of modulus of the reinforced soil K_s was larger when Netlon CE121 was used compared with Netlon CE111.
- The subgrade reaction values for reinforced soil were found to improve by 2 to 5, 3 to 7 , and 4 to 9 times for one layer, two layers, and three layers of reinforcement respectively when compared with those of unreinforced soils. The lower limit reflects the effect of top most reinforcement layer (U is greater than or equal to D), and the upper limit reflects the effect of the case when U is equal D/6. The value of subgrade reaction became steady when U was larger than 0.8D. However, the steady value of subgrade reaction is still larger than that of unreinforced soil.
- The failure criterion in the medium dense reinforced sand has been defined as safe bearing capacity at which settlement is twice the settlement at 60%-75%

of q_s for the case of ($U \leq B/2$), while the reinforced layer at depth of ($U \geq B$), the failure criterion can be defined also near to that of unreinforced and medium sand at 80%-90% percentage of q_s. This amount of reduction in settlement are shown from that the value of $\delta 1 << \delta 2$.

- The safe bearing pressure for footing resting on reinforced soil can be estimated with (Fs=3) from equation (2) after getting (δp) from equation (3); in condition that a plate load test should be achieved.

REFERENCES

Akinmusuru, J.o., and Akinbolade, J. A. (1981). "*Stability of Loaded footing of Reinforced soil*", ASCE Journal of geotechnical engineering, vol. 107, No. GT6, PP. 819-829.

Al-Dobaissi, H. H. (1990). "*Footings on Reinforced Earth Subjected to Impact Loading"*, Master Thesis, University of Baghdad, Iraq.

Ingold, T. S. (1982). "*Reinforce Earth*"; published by Thomas Telford ltd, London, England.

Bassat, R.H. and Last, N.C. (1978). *"Reinforced Earth Below Footing and Embankments"*, ASCE Proc. Conf. Pittsburgh.

Colin JFP Jones (1985). *"Earth Reinforcement and Soil Structures"*. Butterworth and Co. (Publishers)Ltd.

Lambe, T. W., and Whitman, R. U. (1979). "*Soil Mechanics*"; Published by John Wiley and Sons. Inc.

Bieganousky et. ,al 1976;" Uniform placement of Sand",J.of Geo.Eng.Div.,ASCE,Vol.(102),Gt.PP(229-235).

PERFORMANCE IMPROVEMENTS OF ADAPTIVE FIR EQUALIZER USING MODIEFED VERSION OF VSSLMS ALGORITHM

Thamer M.J. Al-anbaky
Department of Electrical and Electronic Engineering, University of Technology

ABSTRACT:
In this paper possible improvements in the performance of adaptive Linear Equalizer (LE) and Decision Feed Back Equalizer (DFE) are reported. A modified Least Mean Square (LMS) algorithm incorporating a recursively adjusted adaptation step size based on rough estimate of the performance surface gradient square is proposed. The first proposed algorithm was called Adjusted Step Size LMS (ASSLMS) which used single adjusted step size for all weight coefficients. The second proposed algorithm was called Distributed Step Size LMS algorithm (DSSLMS). This algorithm (i.e. DSSLMS) will distribute the resultant variable step size in an exponential form among all weights of the adaptive filter such that each weight coefficient has its own step size. These proposed algorithms through computer simulation results shows favorable performance than traditional LMS algorithm and another Variable Step Size LMS (called VSSLMS) algorithm in terms of fast convergence time, less miss-adjustment in steady state, and good tracking ability.

KEYWORDS

Linear Adaptive Equalizer, DFE Equalizer, LMS Adaptive algorithm,

Variable Step Size LMS algorithm.

INTRODUCTION

Adaptive equalizer was widely used in digital communication systems in order to reduce or eliminate the channel distortions like additive noise or intersymbol interference ISI before demodulation at the receiver. The simple structure for adaptive equalizer was the Finite Impulse Response filter (FIR) which can be trained by the Least Mean Square

adaptive algorithm (LMS). This LMS algorithm, which was first proposed by Widrow and Hoff in 1960, is the most widely used adaptive filtering algorithm (Farhang 1999)
This widely applications of LMS algorithm are as a result of its simplicity. Furthermore, it does not require matrix inversion, nor does it require measurements of the pertinent correlation functions (Farhang 1999). But this algorithm suffers from slow convergence since the convergence time of LMS algorithm is inversely proportional to the step size(Widrow and S.Stearns 1985). However if large step size is selected, then fast convergence will be obtained but this selection results in deterioration of the steady state performance (i.e. increased the miss-adjustment (excess error). Also small value of the step size will cause slow convergence but will enhance or decrease the steady state error level(Widrow and S.Stearns 1985).
Numerous modifications of the LMS algorithm have been reported [3, 4, 5, 6, 7, 8, 9 and 10]. In these works, the optimization issue concerning the step size is discussed, and several methods of varying the step size to improve performance of the LMS algorithm especially in time varying environments are proposed. In such environments the step size must be adjusted automatically in order to obtain the following features :

- Adaptive filter must be able to track any change in the system, i.e. to reduce the lag factor which was a decreasing function of the step size.
- Reduce the trade off between the low level of miss-adjustment and fast convergence , i.e. both requirements , must be obtained .

- Reduce the sensitivity of the algorithm to variations in the level of non-stationary, i.e. Low level of miss-adjustment must be obtained when high level of non-stationary occurred.

Optimum step Size can be used but it is undesirable approach since in practical implementation; optimum step size is approximated by trial and error. In addition to the optimum step size cannot track any sudden change in the system environments because it has a fixed value. Some papers used active taps detection techniques with traditional LMS algorithm to enhance the performance of the system. But this technique suffers from overhead operations which is not suitable for hardware implementation. Therefore this paper tries to improve the performance of the LMS algorithm and to get rid of the previous main drawbacks. The proposed algorithms used recursively adjusted adaptation step size based on the performance surface gradient square. The first proposed algorithm was called ASSLMS. Then the obtained step size is distributed among all the weights of the adaptive filter. This proposed algorithm will be refereed to as Distributed Step Size LMS algorithm (DSSLMS) all over in this paper. These proposed algorithms then will be applied to two types of adaptive equalizers which are Linear Equalizer (LE) and Decision Feed Back Equalizers (DFE).

LINEAR AND DECISION FEED BACK EQUALIZER WITH LMS ALGORITHM

Figure (1) shows the classical model of the LE .As shown in this figure there are two modes of operations, namely, the training mode and decision-directed mode (Thamer el al 1997).

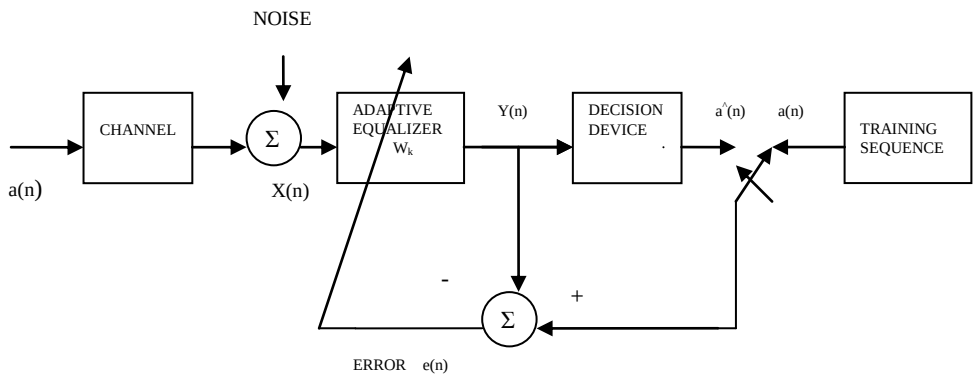


Fig (1) Classical Model of LE

During the training mode, the transmitter generates a data symbol sequence known to the receiver. The receiver therefore, substitutes this known training signal in place of the decision device output. Once an agreed time has elapsed, the decision device output is substituted and the actual data transmission begins. When the training process is completed, the adaptive equalizer is switched to its second mode of operation: the decision-directed mode. In this mode of operation, the error signal is defined by (Thamer el al 1997).

$$e(n) = a^{\wedge}(n) - y(n) \qquad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where y (n) is the equalizer output and $a^{\wedge}(n)$ is the final correct estimate of the transmitted symbol $a(n)$. In figure (2), there is a general structure of the adaptive decision feedback equalizer. This figure shows that DFE equalizer consists of two sections, a feed forward section and a feedback section connected together (Thamer el al 1997)

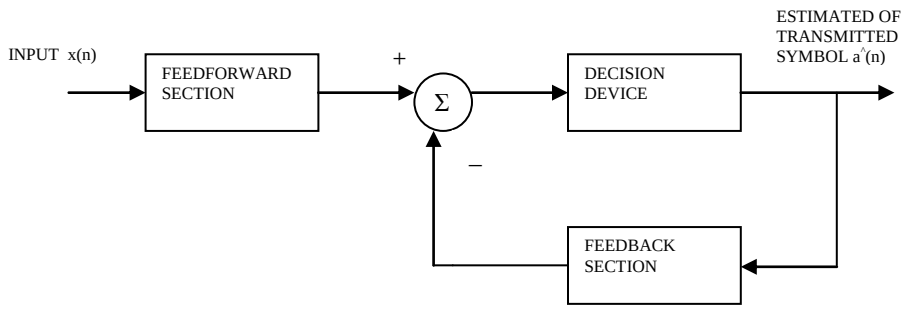


Fig (2) Block Diagram of DFE

The feedback section consists of a tap-delay line filter (such as FIR type) whose taps are spaced at the reciprocal of the signaling rate. The data sequence to be equalized is applied to this section. The feedback section consists of another tapped-delay-line filter whose taps are also spaced at the reciprocal of the signaling rate. The input applied to the feedback section consists of the decision made on previously detected symbols of the input sequence. The function of the feedback section is to subtract out that portion of the ISI produced by previously detected symbols from the estimates of the future samples (Thamer el al 1997). As in the case of the linear adaptive equalizer, the coefficients of the feed forward filter and the feedback filter in a decision-feedback equalizer may be adjusted recursively. Also in the case of a linear equalizer, a training sequence was used to adjust the coefficients of the DFE initially. Upon convergence to optimum coefficients, the system is switched to a decision-directed mode where the decisions at the output of the detector are used in forming the error signal and fed to the feedback filter. This is the adaptive mode of the DFE.

The LMS or Gradient algorithm is used to implement adaptive equalization. It is a stochastic gradient optimization algorithm based on a traditional optimization technique called the Method of Steepest Descent. Let us define the input vector at the equalizer input as $X_n=[x(n), x(n-1),\ldots\ldots,x(n-N-1)]^T$, and the vector of the weight coefficients as $W(n)=[w_0(n)\ w_1(n)\ w_2(n)\ldots w_{N-1}(n)]^T$ of the adaptive filter at an instant n, N is the order of the adaptive FIR filter. Moreover, the signal samples at the equalizer input are of the form:

$$x(n)=\sum_j hn(j)a(n-j)+noise(n) \quad \ldots\ldots\ldots (2)$$

Where a (n) denotes the n-th data transmitted sample, noise (n) is the additive noise with the variance σ^2, and hn (j) is the channel impulse response. The equalizer output at the n-th iteration instant is:

$$y(n)=W^T(n)X(n) \qquad \ldots\ldots\ldots\ldots (3)$$

The output y(n) is used in estimating the transmitted data symbol a(n-D) , with D denoting to the delay. The equalizer output error at the n-th iteration instant in training mode is:

$$e(n) = y(n) - a(n-D) \quad \ldots\ldots\ldots\ldots\ldots \quad (4)$$

The weighting coefficients in the LMS algorithm are obtained from the following expression:

$$W(n+1) = W(n) + \mu e(n) X(n) \quad \ldots\ldots \quad (5)$$

Where μ represents the algorithm fixed step size. The output Mean Square Error (MSE) is(Zens 1989.)

$$\varepsilon(n) = E(e^2(n)) = W^T(n) R W(n) + E(a^2(n)) - 2W^T(n) E(X(n) a(n_0 - D)) \qquad (6)$$

With R=E(X (k) X^T (k)). Where E (.) is expected operation and R is defined as the square matrix. The average output MSE after n-th iteration can be expressed as:

$$\varepsilon_{avr} = \varepsilon_{MIN} + E\left(V^T(n) R V(n)\right) \qquad (7)$$

Where ε_{MIN} is the MSE minimum (Zens1989) ,for optimal weighting coefficients vector $W^*(n)$, and V (k) = W (n)-W^*(n) is the weighting coefficients error vector. In the steady state, the MSE above ε_{MIN} is known as the excess MSE.
As shown in (Widrow and S.Stearns1985), the excess MSE (also called miss-adjustment) for LMS algorithm is given by:

$$\varepsilon_{exess} = \frac{1}{2} \mu \sigma_n^2 tr(R) \qquad \ldots\ldots\ldots\ldots\ldots \quad (8)$$

It may be observed that from (Zens1989) that the excess MSE is due to the gradient noise and is proportional to the step size. The step size must be selected to balance the coefficient goals of the fast convergence (large step size) and small steady state error, i.e. small excess MSE (small step size) (Widrow and S.Stearns1985) .

PROPOSED VARIABLE STEP SIZE LMS ALGORITHMS

ASSLMS Algorithm

The proposed algorithm in this paper is called Adjusted Step Size LMS (ASSLMS) algorithm which used variable step size that will be adjusted according to the square of the gradient of the performance surface (i.e. $e_n X_n)^2$. Using the gradient of the performance surface as a guide to adjust the step size was first developed in(Kang and Johnstone 1992) [9]. Their formula for adjusting the step size was(Simon haykin 1983):-

$$\mu_k = \mu_{\max} .(1 - \exp(-\alpha \| e_k X_k \|) \quad \ldots\ldots\ldots\ldots \quad (9)$$

Where α is a constant called damping factor and $\| \|$ is regular norm vector. This algorithm has practical hardware implementation since equation (9) has exponential factor and require an approximate formula. Another formula to adjust the step size was developed in (Simon haykin 1983). The algorithm used in (Simon haykin 1983) was called Variable Step Size LMS algorithm (VSSLMS) and their formula is:-

$$\mu_n = \alpha\mu_n + \delta e_n^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (10)$$

Where $0\langle\alpha\langle 1$ and $\delta\rangle 0$, then:-

$$\mu_{n+1} = \mu_{\max} \quad \text{if} \quad \mu_{n+1}\rangle\mu_{\max} \quad , \text{ or } \quad \mu_{n+1} = \mu_{\min} \quad \text{if} \quad \mu_{n+1}\langle\mu_{\min} , \quad \ldots\ldots\ldots (11)$$

Otherwise $\mu_{n+1} = \mu_{n+1}$

Where $\mu_{\min}$ is chosen to provide minimum level of miss-adjustment at steady state, and $\mu_{\max}$ ensures the stability of this algorithm (Simon haykin 1983). So this paper proposes to use the square of the gradient estimation $(e_n X_n)^2$ into equation (10) instead of using the square of the error. Then equation (10) will be modified to be as follow(Hulya 2004):-

$$\mu_n = \alpha\mu_n + \delta(e_n . X_n)^2 \qquad \ldots\ldots\ldots (12)$$

This proposed algorithm (ASSLMS) algorithm regard as modified version of the VSSLMS algorithm [12]. Involving the term (X_n) which represents the input signal in the updating step size formula in addition to error factor is favorite choice in order to speed up the estimation and adaptation process.

DSSLMS ALGORITHM

By theory for a stationary channel, the length of the window which tracks the channel is the length of the number of samples. However, channels that have a time- varying nature require a window which must adjust to the recent channel. There are several fundamental considerations that must be understood in the implementation of the LMS algorithm for time- varying channels. First of all, the LMS algorithm uses the most recent channel estimation error into equation (5). This, in turn, means it is severely affected by time-varying channels. Secondly, the most of the impulse response of the channel has decreasing values from higher values to smaller values. Also if all weight coefficients adjusted in equal form will lead to slower the convergence rate of the algorithm. Therefore this paper proposed to adjust the weight coefficients in non uniform manner such that the amount of adjusting the weight coefficients will be in decreasing manner. The only parameter that can perform this function was the variable step size (see equation (5)). This idea can be implemented by applying a sliding window to the obtained step size in equation (12) (Hulya 2004). Such that equation (12) can be distributed among all weight coefficients. There are several types of windows that can be used to do this idea. These can be rectangular, triangular, exponential, and such. In this paper, an exponential sliding window is considered due to its superiority over the other types of windows for the most types of channels. Then the step size that obtained from (12) will be distributed

among the weights of the adaptive filter. Such that each weight coefficient has independent step size $\mu_{i(n)}$ where i=1,2,….,N (order of the FIR filter), and n is the iteration instants. The distribution of the step size μ_{n+1} will be according to the following:

$$\mu_{i(n+1)} = \xi^{i-1}\mu_{(n+1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (13)$$

Where ζ is constant, determine the curvature of exponential distribution of μ_{n+1} among the weights of the adaptive filter. The length of this window is the same as the number of the weight coefficients (i.e. N). Such that each weight will be adjusted independently according to the following:

$$w_{i(n+1)} = w_{i(n)} + \mu_{i(n)} e_n y_{i(n)} \quad \ldots\ldots\ldots\ldots\ldots \quad (14)$$

This proposed algorithm is called Distributed Step Size LMS (DSSLMS) algorithm. This algorithm adapts the filter weights having larger values by amounts more than the weights having smaller values at each iteration. Therefore the weights of coefficients for this proposed algorithm will convergence faster and estimate the inverse impulse response of the channel very quickly.

Simulation Results

<u>LE using Channel A</u>

In this case LE was simulated using different algorithms. The channel used here is called channel A which has a spectral null in the middle frequency region. The impulse response of the channel A is [0.2 -0.15 1.0 0.21 0.03]] . The order of FIR adaptive filter for all simulation was 11 taps and signal to noise ratio was 26 dB, the additive noise was Gaussian noise with mean zero, and variance $\sigma^2 = 0.001$. The training samples were 1000 samples then the adaptive process is switched to decision mode. Figure (3) shows the learning curve for this case with different algorithms. The optimum step size for LMS algorithm was chosen by trial and error to be 0.06.

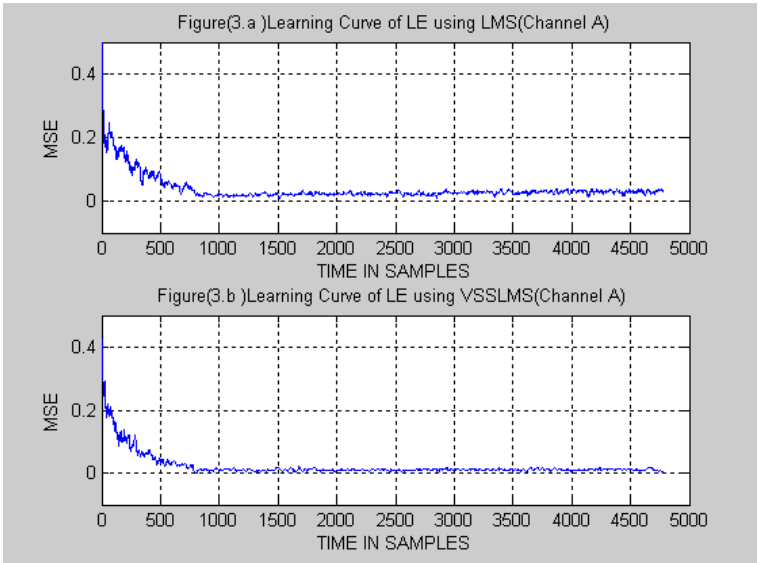


Fig (3)

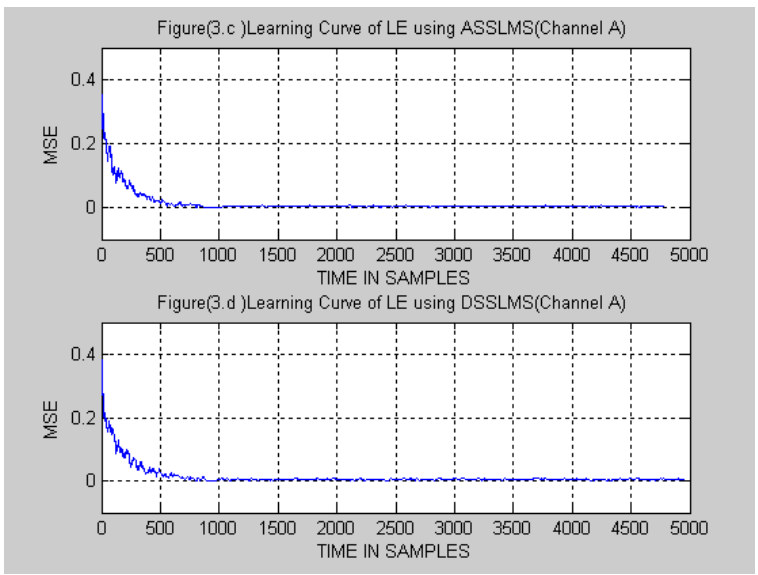


Fig (3) Learning Curves for LE using Channel A

The optimum values of $\mu_{\max}$ and $\mu_{\min}$ was chosen to be 0.06 and 0.0001 respectively for VSSLMS, ASSLMS and DSSLMS algorithms. The values of α and δ was chosen to be 0.97 and 0.001 respectively for all algorithms. The ζ for DSSLMS algorithm was equal to 0.05. As shown in figures (3.c and 3.d) the proposed algorithms have fast convergence time than LMS and VSSLMS algorithms (figures (3.a and (3.b)). Also these proposed algorithms have less miss-adjustment (excess error) in the steady state region than LMS and VSSLMS algorithms.

DFE USING CHANNEL A

In this case DFE is used with the same channel as used previously in LE. The order of feed forward section was 5 and for feedback section was 3. The optimum step size for LMS algorithm was chosen by trial and error to be 0.008. The optimum values of μ_{max} and μ_{min} was chosen to be 0.008 and 0.0000001 respectively for VSSLMS, ASSLMS and DSSLMS algorithms. The values of α and δ was chosen to be 0.97 and 1e-7 respectively for all other algorithms. The ζ for DSSLMS algorithm was equal to 0.05. Figure (4) shows the learning curves for this case.

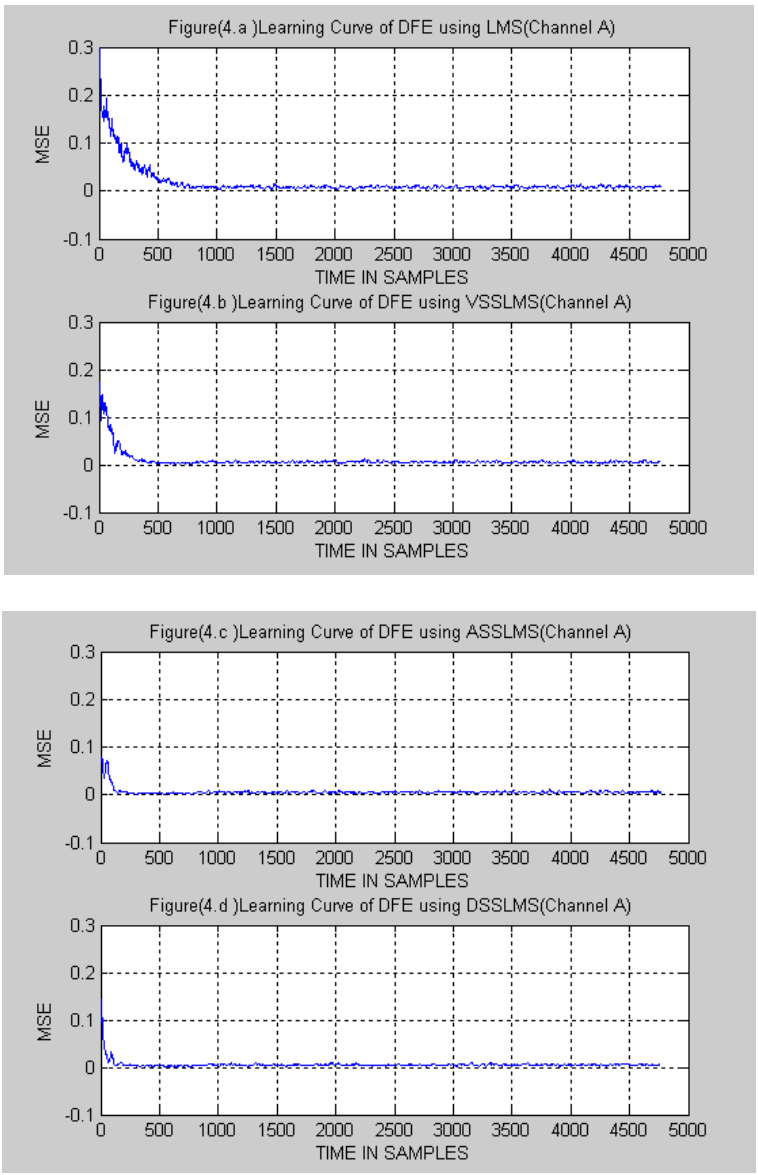


Fig (4) Learning Curves for DFE using Channel A

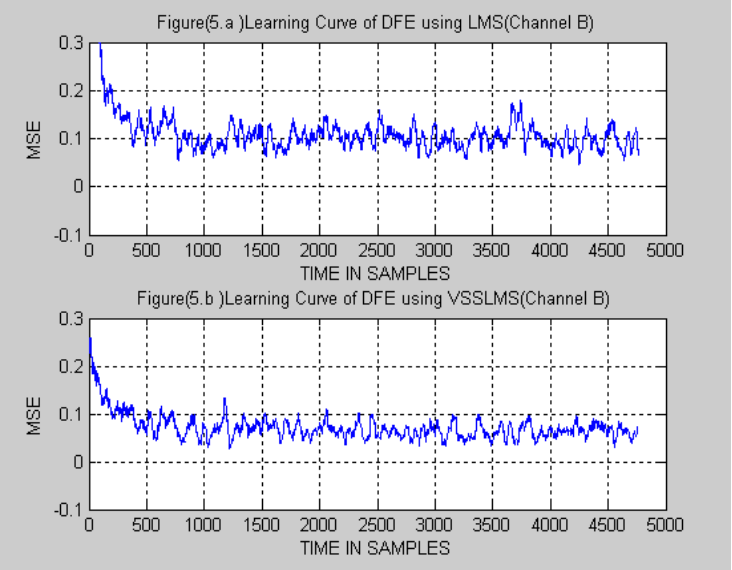
As shown in figure (4) the performance of adaptive equalizer was enhanced using different algorithms with the same channel compared with that obtained using LE type. Also here the performance of the proposed algorithms (figures (4.c and 4.d)) was also better than LMS and VSSLMS algorithms (figures (4.a and 4.b)) in terms of fast convergence and less miss-adjustment in the steady state region.

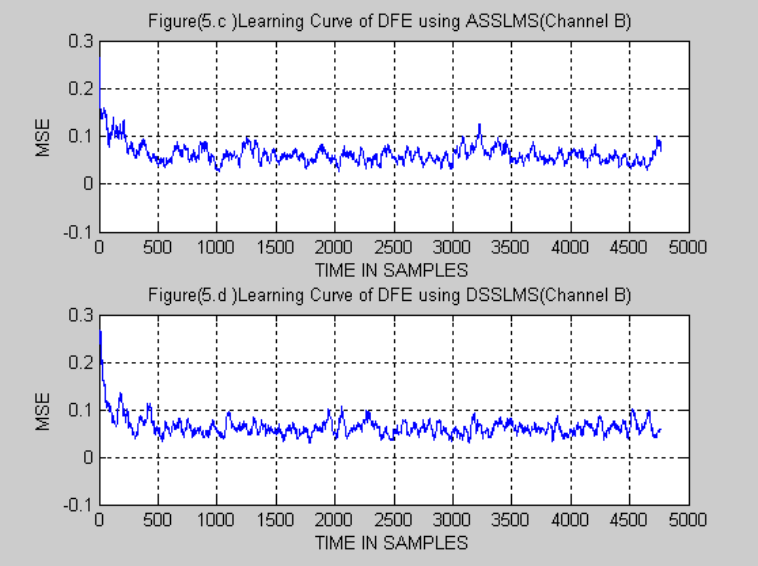
DFE USING CHANNEL B

In this case another channel response is used which has the following impulse response [1 0 1.2 0 0 -0.3]. This channel is called channel B which represent an example of non-minimal phase channel and it is characterized by a multipath that is stronger than the channel A. Figure (5) shows the learning curve of different algorithms using DFE type. The optimum step size for LMS algorithm was chosen by trial and error to be 0.001. The values of $\mu_{\max}$ and $\mu_{\min}$ was chosen to be 0.001 and 0.0000001 respectively for VSSLMS, ASSLMS and DSSLMS algorithms. The values of α and δ was chosen to be 0.97 and 1e-7 respectively for all algorithms. The ζ for DSSLMS algorithm was equal to 0.005. As shown in the figure (5) the good performance of the proposed algorithms (figures (5.c and 5.d) compared of the other algorithms (figures (5.a and 5.b)).

CONCLUSIONS

In this paper possible improvements of the adaptive LE and DFE are proposed by using two modified versions of VSSLMS algorithm. The first one was called ASSLMS algorithm and the second was called DSSLMS algorithm. These proposed algorithms used the square of the gradient estimation $(e_n X_n)^2$ instead of error square alone. Involving the term (X_n) which represents the input received signal in the updating step size formula is favorite choice in order to speed up the estimation and adaptation process. Also using individual step size (DSSLMS algorithm) for each weight coefficients was good idea to speed up and reduce the miss-adjustment in the steady state region. Through simulation results that used two different channels one can see the good performance of the proposed algorithms compared with traditional LMS and VSSLMS algorithm in terms of fast convergence and less miss-adjustment in the steady state region.





Fig(5) Learning Curves for DFE using Channel B

REFERENCES

- B.Farhang Boranjrncy, “Adaptive Filters” John Wileys & Sonc, 1999.

- B. Widrow and S. Stearns,” Adaptive Signal Processing” Prentic-Hall, Inc.1985.

- R.W.Harris, D.M. Chadries , F.A. Bishop “ A variable Step (VS) Adaptive filter algorithm” IEEE Trans. On ASSP-34, No.2, pp.309-316,April 1986. [4] Long Le, Ozgu Ozun, and Phiipp Steurer,” Adaptive Channel Estimation for Sparse Channels”,

submitted for the degree of Bachelor of Eng. (Honors) in the Division of Elect. And Electronic Eng. , university of Queensland, 2002..

- Charles Q. Hoang," LMS Modeling and Estimation of Non-stationary Telecommunication Channels", submitted for the degree of Bachelor of Eng. (Honors) in the Division of Elect. And Electronic Eng. , university of Queensland, 2000.

- J.J. Chen, R.R. Priemer " An in Inequality by which to adjust the LMS algorithm Step Size" IEEE Trans. On COMM-43,No.2/3/4,pp1477-1483,Feb.1995

- Long Le, Ozgu Ozun, and Phiipp Steurer," Adaptive Channel Estimation for Sparse Channels", submitted to the Division of Elect. And Electronic Eng. , university of Queensland, 2002..

- Bozo K. ,Zdravko U. , and Ljubisa S. ,"Adaptive Equalizer with Zero-Noise Constrained LMS algorithm ", In Elec. Eng. of FACTA University, vol.16, April 2003, 1-xx.

- S.K., G. Zeng." A new Convergence Factor for Adaptive Filters" IEEE Trans. On CAS-36,No.7,pp.1011-1012, July 1989.

- R.H.Kang, E.W.Johnstone,"A variable Step Size LMS algorithm" IEEE Trans. On Signal Processing, Vol.40, No.7, pp.1633-1642, July 1992.

- Simon Haykin ," Communication Systems" John Wiley & Sons, Inc., New York, 1983.
- Thamer M. Al-Anbaky , Manal J. Al-Akindy, and Aide K. Al-Samarrie "Performance Improvement of Adaptive FIR Filters Using Adjusted Step Size LMS Alogorithm" In Seventh International Conference on HF Radio Systems and Techniques, Nottingham, UK, 7-10 July 1997.

- "ADAPTIVE EQUALIZER EXPERIMENTS" **by** Hulya Seferoglu Faculty of Engineering and Natural Science, Sabanci University, Istanbul, Turkey Spring 2004.

OPTOELECTRONIC IMPLEMENTATION OF ARTIFICIAL NEURAL NETWORK: PERCEPTRON LEARNING RULE AND M-CATEGORY CLASSIFIER

Asst. Prof. Dr. M. S. Abdul-Wahab Asst. Prof. Dr. Hanan A. Reda Akkar
O. Q. J. Al-Thahab
Department of Electrical And Electronic Engineering
University of Technology
Baghdad – Iraq

ABSTRACT

Single neuron perceptron is designed as a classifier of two different classes using the hard-limiter activation function (i.e. in the absence of light, and presence of light). An example is designed and tested so that the proposed circuit learned different categories and then used as a classifier for two different classes because of the use of single neuron. Additional electronic circuits were used for computation processes. The Computer simulation results indicate stable solution that compares with theoretical results.

Single layer perceptron M-category classifier is designed as a classifier for more than two classes. An example is designed and tested for the verification. The example learns after (5) iterations. Computer simulation results indicate stable solution that compares favorably with theoretical results.

الخلاصة

تم اعتماد نموذج برسبترون كأحد النماذج التي تطبق بوجود المشرف، وتم فحص مثال على الدائرة المقترحة، حيث تم تعليم الدائرة أنماط مختلفة ومن ثم استخدامها كمصنِّف لنمطين فقط بسبب استخدام خلية عصبية واحدة. لقد استُخدمت دوائر إلكترونية أخرى مساعدة لتحقيق العمليات الحسابية التي نحتاجها في تمثيل هذا النموذج، وقد أظهرت الدائرة المقترحة نتائج مقاربة للنتائج النظرية من خلال محاكاة الدائرة المقترحة حاسوبياً.

كذلك تم تصميم نموذج برسبترون لتصنيف الأنماط المتعددة من أجل تصنيف عدد أكبر من الأنماط (أكثر من اثنين)، وقد تم فحص مثال عليها وكانت النتائج مقاربة للنتائج النظرية، كما وتم استخدام دالة المحدد كدالة فعالة في الدائرة المقترحة.

INTRODUCTION

Learning theory develops models from data in an inductive framework. It is, therefore, no surprise that one of the critical issues of learning is generalization. But before generalization, the machine must learn from data [Principe 2000]. Learning can be viewed as maximizing the likelihood of observed data under the generative model, which is mathematically equivalent to discovering efficient ways of coding the sensory data [Chahramani 1999]. While neural networks can and are implemented entirely with electronic hardware, optoelectronics have a clear advantage in the task of interconnecting neurons.

Unlike electronic signal, light beams do not interact with one another, permitting many connections to be made in the same space. Since photons do not interact with one another, it is difficult to implement directly the processing operation of the neuron with optical devices, thus many systems combine optical synaptic interconnections with optoelectronic neurons [Wilamowski 2000].

In this paper, a variety of issues are discussed that have impact on the development of optoelectronic technology as applied to hardware implementation of neural networks with enhanced capabilities. Optoelectronics has the potential for the implementation of neural networks with large numbers of neurons (10^5 to 10^6) and high connectivity (approximately 10^{10} analog weighted interconnections) in one module. The approach taken here is to use electronics to implement the internal functions of each neuron unit, and to use optics to implement the connections, weight, and input and output. With this technique, most of the area of a two-dimensional "chip" can be used to implement the neuron units themselves [Keller 1993]. Analog optoelectronic hardware implementation of neural nets has been the focus of attention for several reasons. Primary among these is that the optoelectronic approach combines the best of two worlds: the massive interconnectivity and parallelism of optics, the flexibility, high gain, and decision making capability (non linearity) offered by electronics [Keller 1992].

OPTOELECTRONIC NEURAL NETWORK

Optical implementation of neural networks is promising for a variety of reasons. First of all, light offers the fastest possible communication channel, not requiring physical limiting conductors. Secondly, increasing light beams do not noticeably interfere with each other. This means that the large number of interconnections of artificial neural network can be optically implemented in a compact way of truly parallel implementation of large neural networks. Optical implementations of multilayer neural network perform nonlinear thresholding which is an essential constituent of all neural network models, and hence involve conversion of optical signals to electronic ones and vice versa. In order to avoid this conversion and to progress to all – optical forward propagation in multilayer neural networks, the use of optical activation function is essential [Al-Buhrezi 2001].

Desirable features for the field of optical neural networks are the ability of performing subtraction and an ideal optical linearity, which is spatially uniform in its response. In general, spatial non - uniformities are measured as a variation in the read out intensities.

The fan – out optics and other optical elements may have non-ideal behavior. These non-uniformities are expected to be compensated to a considerable extent in an adaptive optical neural network, the weights of which are updated during the training of the actual optical system.

Generally, in optoelectronic neural system each neuron is composed of an input summing port, nonlinear transfer device, and an output port. Differential pair of detectors is operated as the input to the neuron; signals with positive (excitatory) weights arrive at one detector and signals with negative (inhibitory) weights arrive at the other detector. These detectors sum the intensity of each optical signal arriving at the neuron. The neuron's activation function is electronically applied to the detected signal to produce an output signal. The output signal drives either an optical source or pair of sources [Chew 1993].

PERCEPTRON LEARNING RULE

For the perceptron learning rule, the learning signal is the difference between the desired and the actual neuron's output. Thus, learning is supervised and the learning signal is equal to:-

$$r = di - oi \qquad \dots (1)$$

where $oi = f(w_i^t x)$, and (di) is the desired response as shown in **Fig.(1)**. Weight adjustment in this method, (Δwi), and (Δwij), is obtained as follows:-

$$\Delta wi = c[di - f(w_i^t x)]x \qquad \dots (2a)$$

$$\Delta wij = c[di - f(w_i^t x)]xj , \qquad \dots (2b)$$

for j=1, 2, …, n and t:- iteration

Under this rule, weights are adjusted if and only if (oi) is incorrect. Error as a necessary condition of learning is inherently included in this training rule. Obviously, since the desired response is either (1) or (-1), the weight adjustment (2a) reduces to:-

$$\Delta wi = \pm 2cx \qquad \dots (3)$$

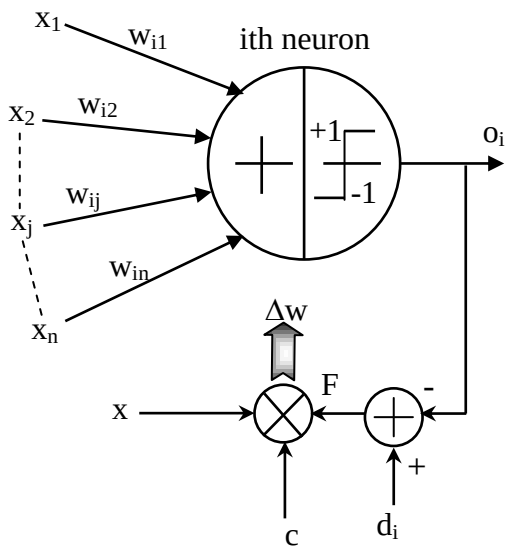


Fig. (1) Perceptron learning rule.

EQUIPMENT NEEDS FOR THE DESIGN OF PERCEPTRON LEARNING RULE

Even though neural networks are primarily implemented in software, their good approximation properties make them attractive in hardware [Moerland 1996], as seen from the previous section (i.e. from perceptron model). First the multiplication of the input signal by the stored weight must be performed. A simple multiplier can achieve this, and it must be four-quadrant multiplier, use of the four quadrant analog Gilbert multiplier satisfies this purpose. After this stage, a summing device is used. It can be seen that the summer is either for voltage or for currents, as seen from **Fig.(1).**

SUMMER-DRIVER CIRCUIT

The summing circuit is shown in **Fig.(2),** it can be seen that if the summation is grater than (zero) then the output will be in logic (0) because the transistor reaches the saturation region or the output will be logic (1) when the value of summation is less then (zero) because the transistor falls onto the cut off region.

This circuit will be used to drive the light emitting diode as shown in **Fig.(3)**. When the transistor turns (ON) and if (Ic) is greater than the turn (ON) current of the (LED), the (LED) will light up, and this can be represented as the (ON) state or logic (1). Else if (Ic) is less than the turn (ON) current of the (LED), the (LED) will not light up, so this case can be considered as logic (0). The turn (ON) current for the (LED) used in this circuit is (10 mA).

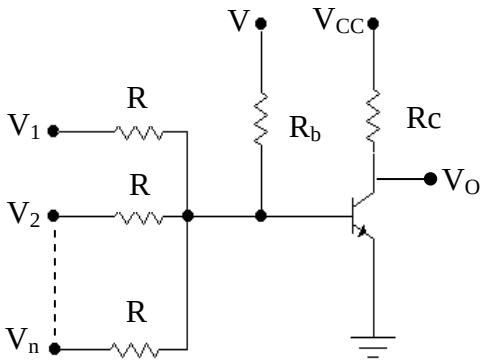


Fig. (2) The current summing circuit

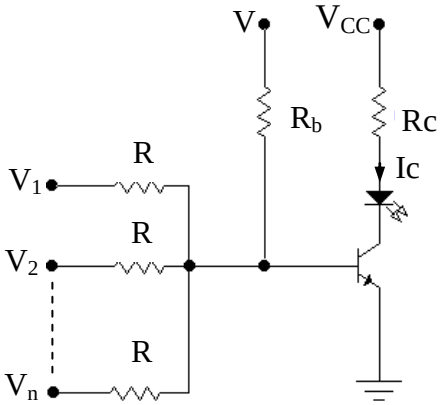


Fig.(3) The sum-driving circuit

FOUR QUADRANT ANALOG GILBERT MULTIPLIER

A Four Quadrant Analog multiplier is a device in which the output voltage is directly proportional to the product of the two input voltages regardless of the polarity of the inputs thus:-

$$V_O = k\, V_1 V_2 \qquad \ldots (4)$$

where (V_1) and (V_2) can be either positive or negative.

The basic Four Quadrant Analog multiplier is Gilbert multiplier as shown in **Fig.(4),** having:-

$$V_{out} = Rc\ Ix\ \tanh\left(\frac{V_1}{2Vt}\right)\ \tanh\left(\frac{V_2}{2Vt}\right) \qquad \ldots (5)$$

If (V_1) and (V_2) << V_t then equation (4) can be written as:-

$$V_{out} = Rc\ \frac{V_1 V_2}{4Vt^2} \qquad \ldots (6)$$

Where V_t is the threshold voltage.

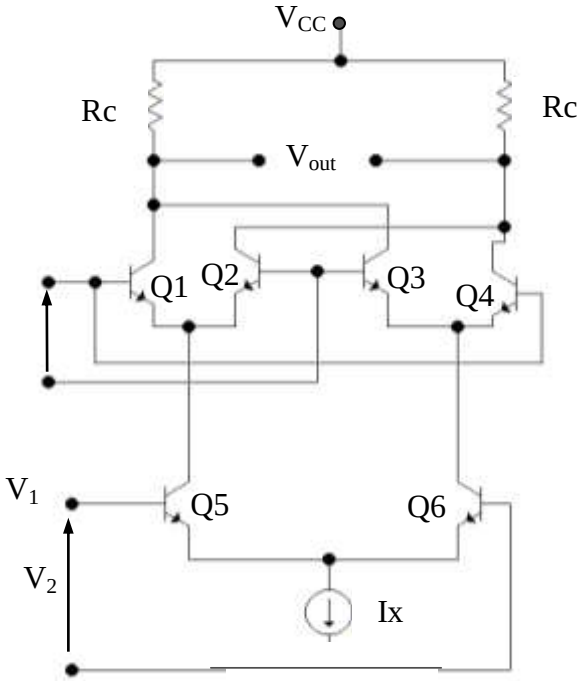


Fig. (4) Gilbert multiplier.

The current source as shown in **Fig. (5)**, illustrate that:-

$$Ix= \frac{1}{R_3}\left(\frac{V_{EE}R_2}{R_1+R_2}+\frac{V_D R_1}{R_1+R_2} - V_{BE7}\right) \quad \ldots (7)$$

Since this current is independent of the signal voltages (V_1) and (V_2), then (Q7) acts a supply of constant current to the Gilbert multiplier.

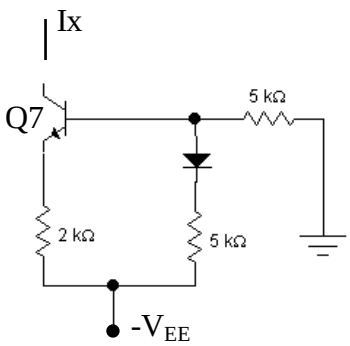


Fig. (5) The current source circuit

The CMOS Inverter

The (CMOS) inverter circuit is shown in **Fig.(6)**, which consists of complementary (NMOS) and (PMOS) transistors, and having (V_{SS}=-V_{DD}), Vthn= -Vthp (Vthn=1.4 V and Vthp= -1.4 V). If we want to use this circuit to be in logic mode then V_{SS} is set to zero.

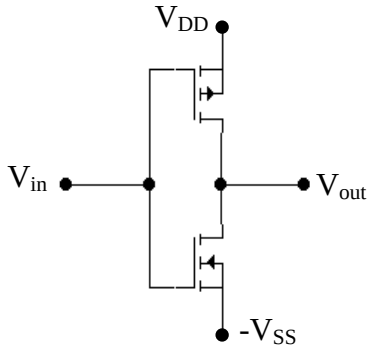


Fig.(6) The CMOS inverter circuit.

THE CMOS ANALOG SWITCH

Connecting an (NMOS) device in parallel with a (PMOS) forms a (CMOS) analog switch also called bilateral transmission gate. This combination acts like two parallel switches, both of which are simultaneously closed or simultaneously open. Actually when the switches are closed, both conduct small positive and negative signals, but only one conducts large positive signals, while other conducts large negative signals. **Fig.(7)** shows how the devices are connected.

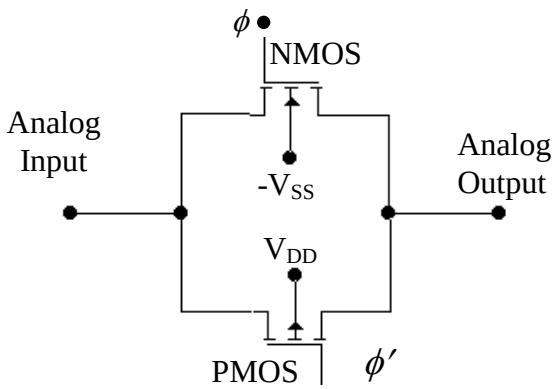


Fig. (7) The CMOS analog switch.

PROPOSED DESIGN OF SINGLE NEURON PERCEPTRON LEARNING RULE

The main problem in all learning rule designs is the storage element (i.e. where the updating weight is stored), and it can be taken as a capacitor. The synaptic weight can thus be stored as a charge on the capacitor (storage capacitor).

Fig.(8) illustrates the block diagram of the proposed circuit for the perceptron learning rule for (n) inputs. It includes elements of all above mentioned circuits. This circuit needs two storage capacitors: the first one, (C1) is used to store the updated weight w(t+1) then transfers its charge to the second capacitor (C2) which is used to store the weight (w(t)) as charge (C1=1 μF. and C2=10 μF.).

The photodiode used in the design of this circuit supplies a current (Iph=100 μA). Ones sees that when the photodiode is illuminated by an optical signal the output of the drive circuit will be (-1) in order to achieve the subtraction of the neuron output from the desired signal, and when there is no optical signal then (Iph= 0), the neuron output will be (1) for the

same reason mentioned above. A simple subtraction circuit can be used to achieve this purpose.

The initialization of weights is stored on the storage capacitor before the learning operation takes place. The proposal presented above has been implemented with the aid of the software package (EWB) program, and verified by (MATLAB) programs.

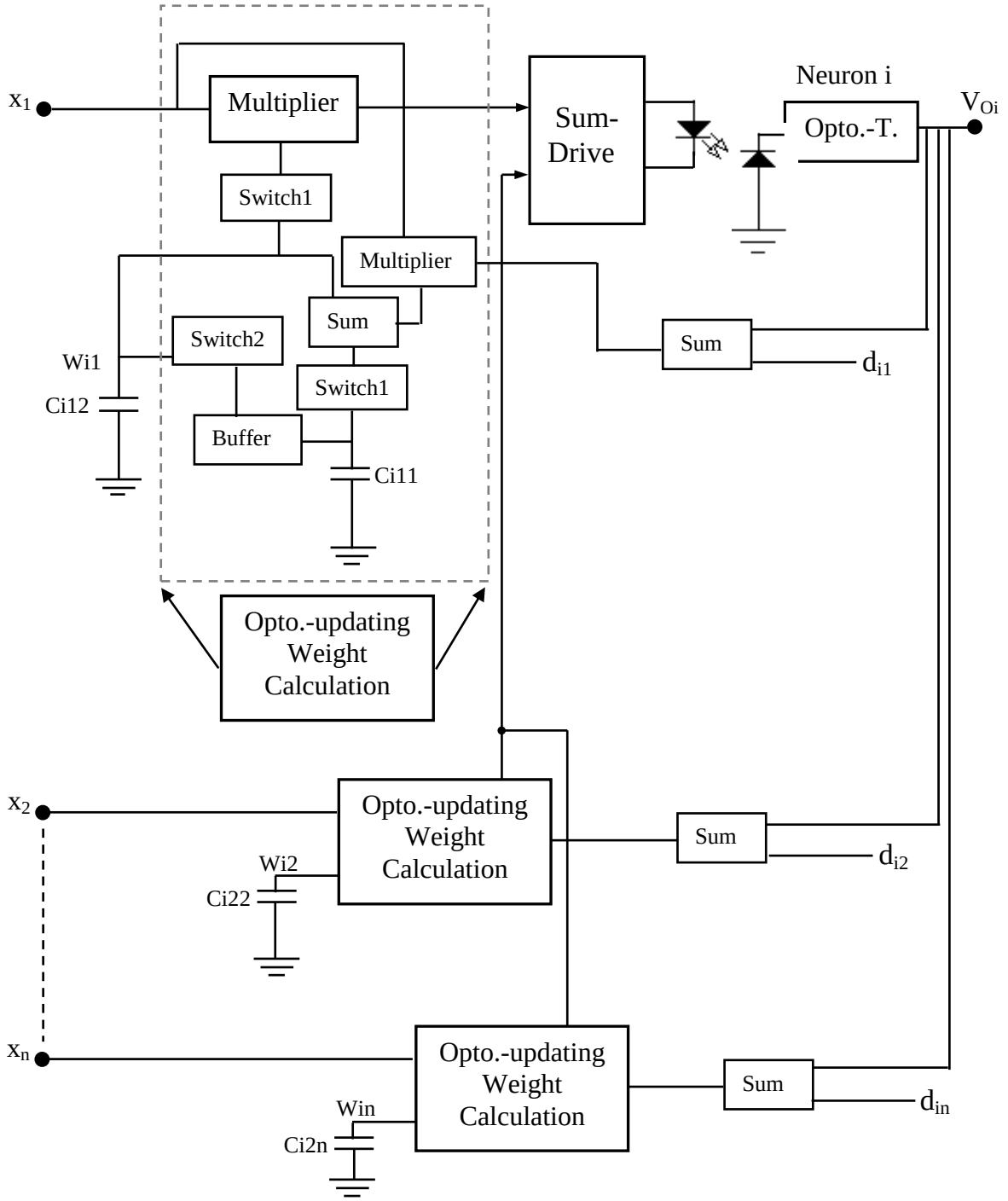


Fig.(8) The block diagram of the proposed implementation circuit for perceptron learning rule.

EXAMPLE 1

A single neuron can be designed to classify two different input patterns. **Fig. (9)** shows an example of two patterns (M=2 and number of nodes n=12), which are the characters (O, j) entered to the proposed circuit of the perceptron rule with the desire (1, -1) respectively, The circuit will learn until it reaches the final weight when w(t+1)=w(t).

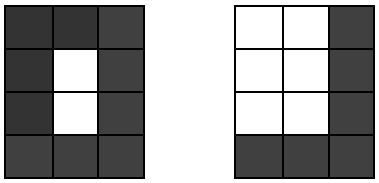


Fig. (9) Two different input patterns.

The results are illustrated in **table (1)** obtained from the simulation results of the (MATLAB) computer programs for (example 1). It can be seen that the learning process stopped after three iterations. As an examination, two distorted patterns are examinated. The vector of the distorted pattern shown in **table (1)**, has the shape seen in **Fig. (10)** together with the correct pattern. The proposed circuit block diagram of this example is shown in **Fig.(11).**

Table (1) Simulation result for example 1.

Pattern 1	Pattern 2	Initial Weight W^0	Updating Weight W^3	Distorted Input	
				Pattern 1	Pattern 2
1	-1	-.5	-2.5000	1	1
-1	1	-1	-3.0000	-1	1
-1	-1	.3	-1.7000	-1	1
1	-1	-.5	-2.5000	1	-1
-1	1	.1	-1.9000	-1	-1
-1	-1	0	-2.0000	1	1
1	-1	.3	-1.7000	-1	-1
-1	1	.1	-1.9000	-1	-1
-1	-1	-.4	-2.4000	1	1
1	-1	-.2	-2.2000	1	1
1	1	-.3	-2.3000	1	1
1	-1	1	-1.0000	-1	1

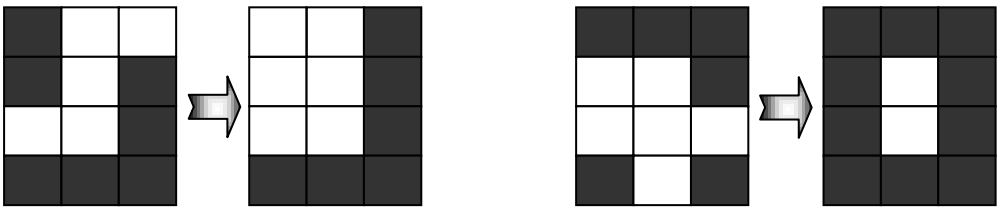


Fig. (10) Distorted pattern examination

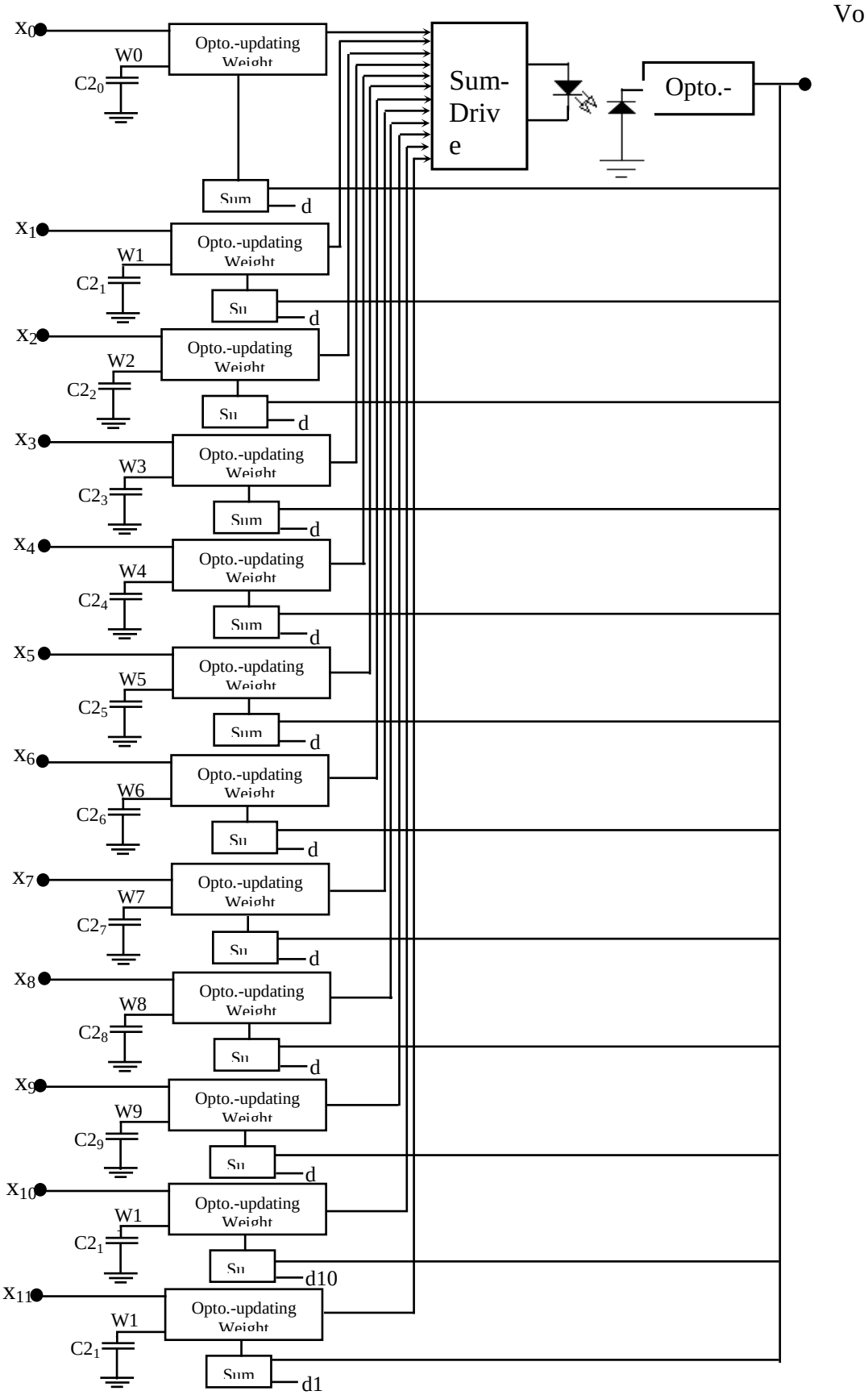


Fig. (11) Optoelectronic implementation of (3*4) input pattern

MULTICATEGORY SINGLE LAYER PERCEPTRON NETWORK

In this section an attempt to apply the error-correcting algorithm to the task of multicategory classification is made. The assumption needed is that classes are linearly pair wise separable, or that each class is linearly separable from other class. This assumption is equivalent to the fact there exist (M) linear discriminate functions such that:-

$$g_i(x) > g_j(x) \qquad \text{for } i,j = 1, 2, \ldots, M \;,\; i \neq j \qquad \ldots (8)$$

where g(x) is the discriminate function and can be considered as an equation depending on the stored weight. Since the inputs here are not only (x), but is added to it a constant value called (bias) which always has the value (-1). This bias element makes the learning process faster so the discriminate function according to this input is given by:-

$$y = \begin{bmatrix} x \\ -1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ -1 \end{bmatrix} \begin{matrix} y_1 \\ y_2 \\ \\ \\ y_{n+1} \end{matrix}$$

$$g(x) = w_1x_1 + w_2x_2 + \ldots + w_nx_n - w_{n+1} \qquad \ldots (9)$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ w_{n+1} \end{bmatrix}$$

In this paper the focus is (M) discrete perceptron. The network generated in this way is comprised of (M)-discrete perceptrons as shown in **Fig.(12).** The output will depend on the discriminate function g(x), so if g1(x) > gj(x) where j=1, 2, …, M, then (O_1=1) and (O_2, O_3, …, O_M) is equal to (-1). This theory exists when using a hard-limiter as an activation function shown in **Fig.(12).** It is called Threshold Logic Unit (TLU #M). This should indicate category (1) input. So the classifier using (M) individual (TLU) elements can be obtained. For this classifier the (k'th) (TLU) response of (1) is indicative of class (k) and all other (TLUs) respond with (-1).

The weight adjustment during the (k'th) step for this network is as follows:-

$$w_i^{k+1} = w_i^k + \frac{c^k}{2}(d_i^k - O_i^k)\, y^k \qquad \text{for } i=1, 2, \ldots, M \qquad \ldots (10)$$

where (di) and (Oi) are the desired and the actual response of the (i'th) discrete perceptron respectively. The desired response for the training pattern of the (i'th) category is:-

$$d_i = 1 \;,\; d_j = -1 \qquad \text{for } j=1, 2, \ldots, M. \quad i \neq j \qquad \ldots (11)$$

Example

Assume having three classes with random initial weights as shown in **table (3),** and the augment pattern are presented in the sequence (y_1, y_2, y_3, y_1, y_2, …) as shown in **Fig.(13a).**

Table (3) Input classes and initial weights.

Class 1	*Class 2*	*Class 3*	w_1^0	w_2^0	w_3^0
10	2	5	1	2	4
6	-4	-2	-1	-1	3
-1	-1	-1	0	2	0

Step 1: y_1 is input**: -**

O_1=1, O_2=1*, O_3=1*

Since the only incorrect response is provided by (TLU #2, 3), so for (c=1):

$$w_1^1 = w_1^0, \quad w_2^1 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 10 \\ 6 \\ -1 \end{bmatrix} = \begin{bmatrix} -8 \\ -7 \\ 3 \end{bmatrix}$$

$$w_3^1 = \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix} - \begin{bmatrix} 10 \\ 6 \\ -1 \end{bmatrix} = \begin{bmatrix} -6 \\ -3 \\ 1 \end{bmatrix}$$

Step 2: y_2 is input**:-**

O_1=1*, O_2=1, O_3=-1

The weight updates are:-

$$w_2^2 = w_2^1, \quad w_3^2 = w_3^1,$$

$$w_1^2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ -4 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$$

Step 3: y_3 is input**:-**

O_1=-1, O_2=-1, O_3=-1*

The weight updates are

$$w_2^3 = w_2^2 \quad w_1^3 = w_1^2$$

$$w_3^3 = \begin{bmatrix} -6 \\ -3 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \\ 0 \end{bmatrix}$$

This terminates the first learning cycle and the third step of weight adjustment. The only adjusted weights as from now will be those of the third perceptron. The outcome of the subsequent training is

$$w_3^{77} = \begin{bmatrix} 3 \\ -3 \\ 20 \end{bmatrix}$$

The three-perceptron network obtained as a result of the training is shown in **Fig.(13b).** It perform the following classification:

$O_1 = f(-x_1 + 3x_2 - 1)$

$O_2 = f(-8x_1 - 7x_2 - 3)$

$O_3 = f(3x_1 - 3x_2 - 20)$

The resulting decision surfaces are shown in **Fig.(14).** It can be seen that the three – perceptron classifier produces three decision surfaces (shown as lines here) and they are:

$-x_1 + 3x_2 - 1 = 0$

$-8x_1 - 7x_2 - 3 = 0$

$3x_1 - 3x_2 - 20 = 0$

The lines are shown in **Fig.(14)** along with their corresponding normal vectors directed toward their positive sides. It can be seen that there are several indecision regions thus patterns in shaded areas are not assigned any reasonable classification. One such patterns may be (Q). The corresponding linear discriminate functions are shown in **Fig.(15).**

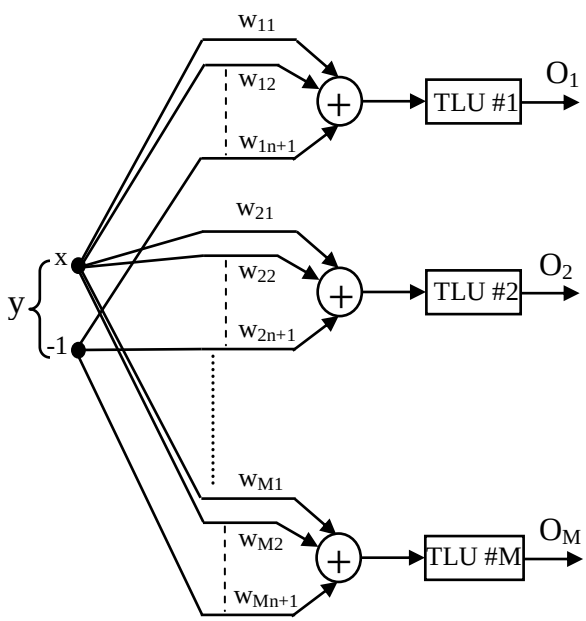
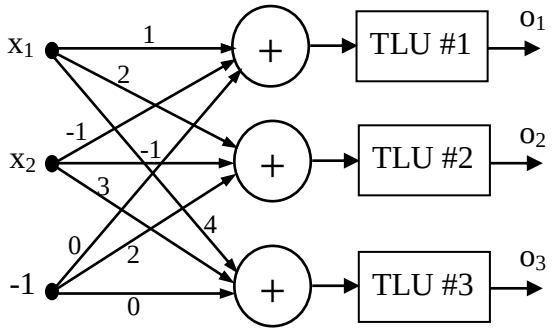
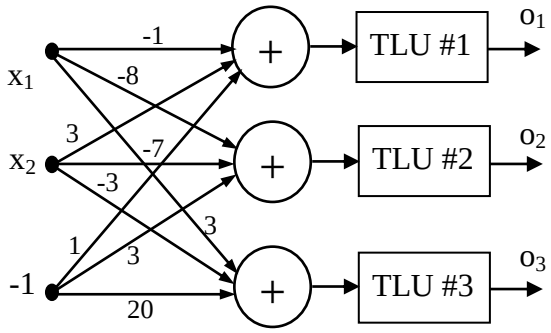


Fig.(12) M-category linear classifier using M-discrete perceptron



(a)



(b)

Fig. (13) Three-class classifier(a) three perceptron untrained classifier. (b) three perceptron trained classifier.

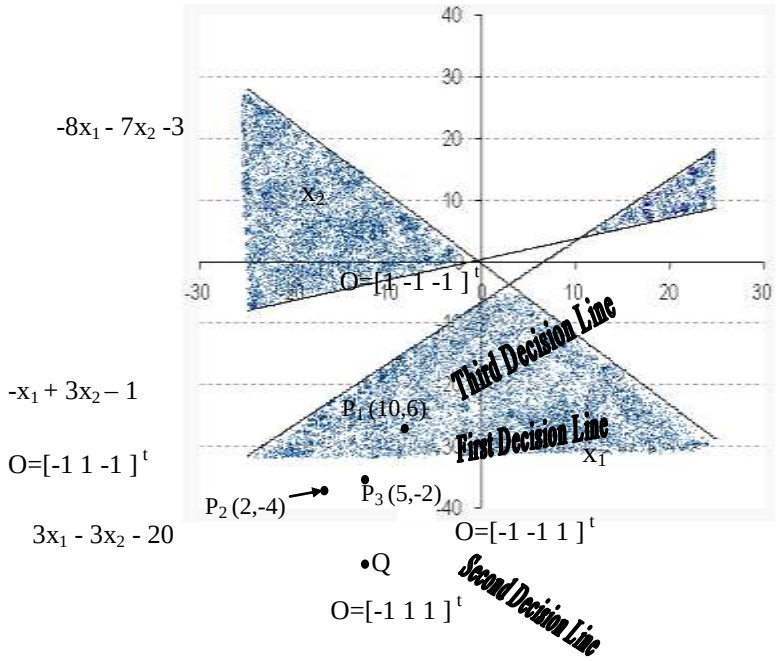


Fig.(14) Decision regions.

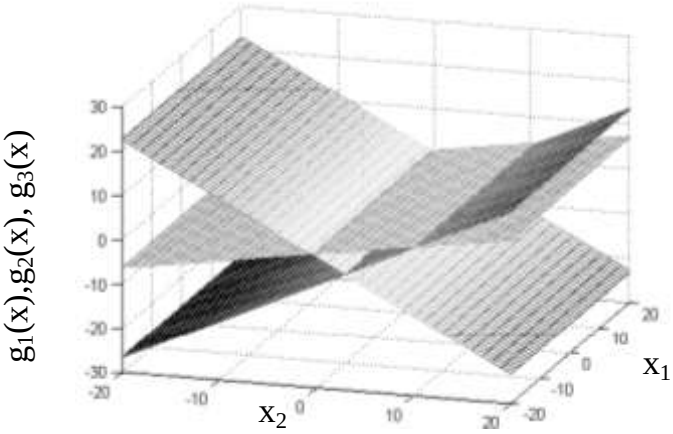


Fig. (15) Discriminate functions.

PROPOSED OPTOELECTRONIC CIRCUIT FOR MULTICATEGORY SINGLE LAYER PERCEPTRON

The sub-circuits needed for the designs of M-category single layer perceptron are similar to those used for the design of single neuron perceptron learning rule. The M-category depends on the same model of the single neuron perceptron except the learning constant in M-category perceptron is half that of the single layer perceptron model used. This can be achieved by multiplying the subtraction between the output and the desired by (0.5), by using the same operational-amplifier circuit.

The block diagram of the proposed circuit for (n) input and (M) category is shown in **Fig. (16).**

EXAMPLE 3

To examine the proposed circuit, three classes are applied to it as follows: the three classes are three characters (T, F, I), as shown in **Fig.(17),** are entered to the circuit. The input signals and the final training weights are tabulated in **table (4)**.

The results were obtained by using (MATLAB) programs for computer simulation as shown in **table (4).** These results were obtained when the desired input was of the form:

$$d = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$

entered into the circuit.

The learning process stopped when the output signal was equal to the desired even though the input classes contained training because the training process is not affected by the updating of weight.

The problem in this classifier model is the value of the bias so that the patterns that are used to classify it must have the value of (-1) in its last term, but this problem can be solved by increasing the number of pixels used to divide the input pattern or adding an additional column so that it takes the value of (-1).

The entire components used in the design of the proposed circuits were designed using (Electronic Work Bench) program in computer simulation. As a test, three different classes were entered to the proposed circuit for their recognition them. The results are shown in **Fig. (18).**

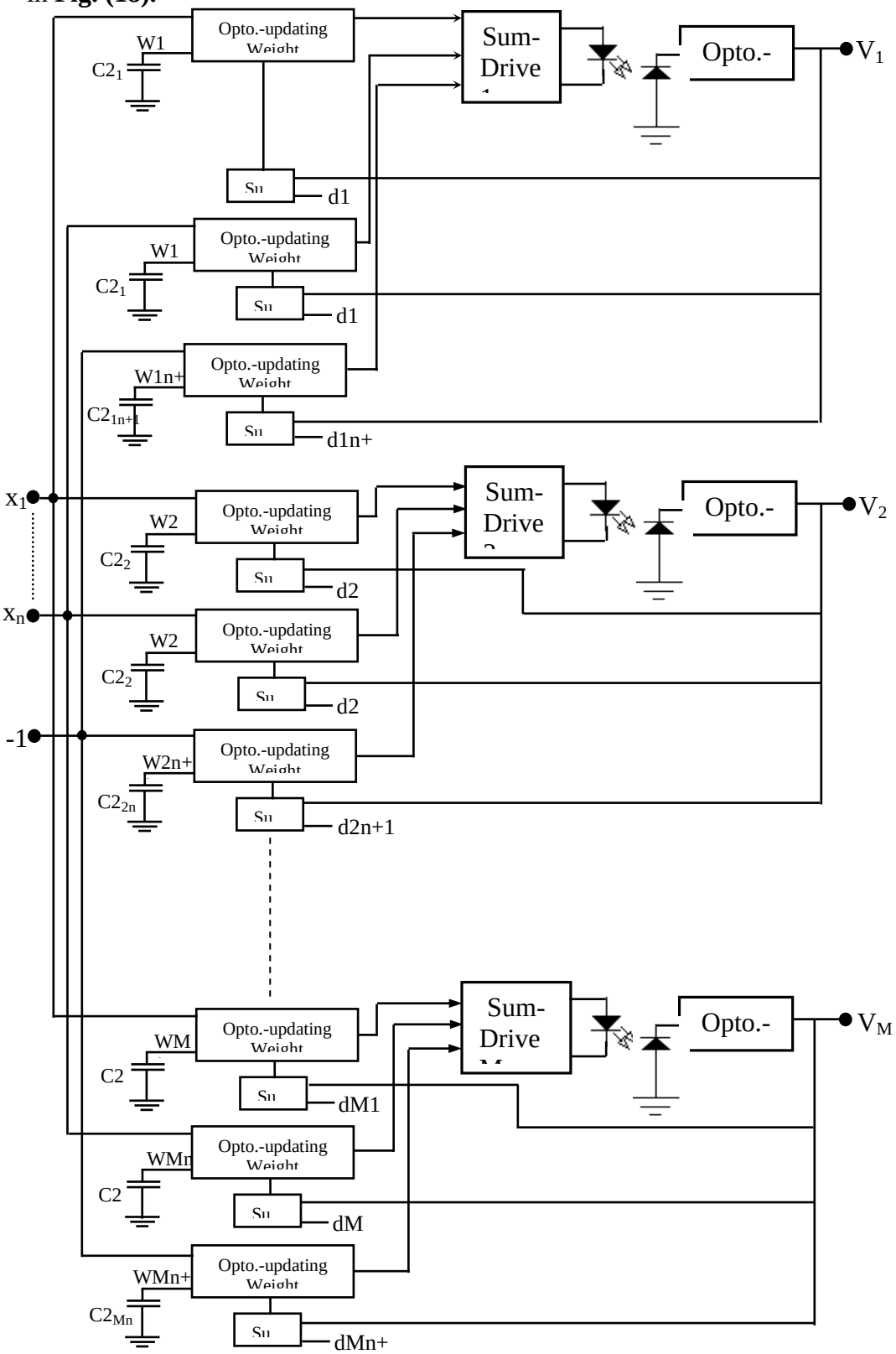


Fig. (16) M-category classifier for n-input signals block diagram

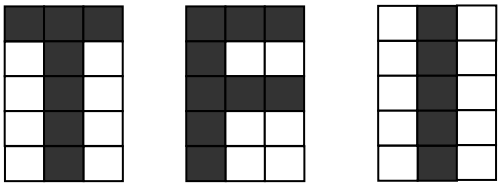


Fig. (17) Three input characters.

Table (4) Input signals, initial weight, and final training.

Class 1	Class 2	Class 3	w_1^0	w_2^0	w_3^0	w_1^5	w_2^5	w_3^5
1	1	-1	1	.5	1	4	-0.5	0
1	1	1	.5	1	-1	-0.5	0	-2
1	1	-1	0	-1	0	3	-2	-1
-1	1	-1	-1	0	.5	-2	1	-0.5
1	-1	1	-1	-.8	0	0	-1.8	1
-1	-1	-1	0	-1	.3	1		1.3
-1	1	-1	.4	.2	0	-0.6	1.2	-1
1	1	1	2	.6	-.2	1	-0.4	-1.2
-1	1	-1	.1	-.1	-.2	-0.9 -	0.9	-1.2
-1	1	-1	.4	.9	-.2	0.6	1.9	-1.2
1	-1	1	.2	.7	-.5	1.2	-0.3	0.5
-1	-1	-1	-.3	-.6	.7	0.7	0.4	1.7
-1	1	-1	1	-.1	.35	0	0.9	-0.65
1	-1	1	.5	-.25	-1	1.5	-1.25	0
-1	-1	-1	.35	1	.6	1.35	2	1.6

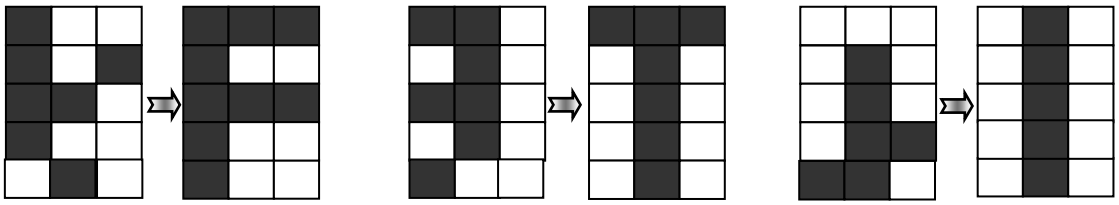


Fig. (18) Three different distorted classes and their correct class.

DISCUSSION

Neural networks and classical optical processing can be combined together to create optical neural networks, which offer great and fundamental advantages over electronic neural networks in various well- defined cases.

The fundamental advantage of optics is solving the massive interconnection between processors required in most neural network models. Since in electrical systems the electrical signals must travel on physical wire, then the massive interconnectivity is very difficult to be

achieved. It requires a large area and careful design to minimize the interference between the physical wires and crosstalk. The perceptron learning rule is an example of supervised learning rule. Single neuron perceptron can be used to classify only two different classes, so the output may be taken as (1) for excitatory case and (-1) for inhibitory case. This is achieved by the LED driving circuit and it can be considered as hard limiter activation function, and the desired sums with output.

The optoelectronic representation of M-category single layer perceptron was realized. It depends on the single neuron perceptron model, it can be used as a classifier with its output depending on the desire that the output must reach it, such that the number of neurons is equal to the number of classes. In this design another element, which is very important in the classification process, is seen, it is the bias that speeds the learning process.

CONCLUSION

The availability of the desired speeds the learning process and lowers the number of iterations needed for reaching the final weight. This can be seen from the example taken on the proposed design of single neuron perceptron by using the same initial weight used for Hebbian learning rule so that the updating weight reaches the final weight after (3) iterations only. The designed M-category classifier can be used to classify the input pattern into one of a very large number of categories due to the large number of synapses capacity per neuron.

REFERENCES

B.M Wilamowski, J. Binfet, and M. O. Kaynak, **VLSI Implementations of Neural Networks**, www.com/ Goggle/ Neural Networks VLSI Implementation, 2000.

C. P. Chew, R. W. Newcomb, and J. D. Yuh, **VLSI Circuits For Optoelectronic Neural Network Weight Setting**, Microsystems Laboratory, Electrical Engineering Department, University of Maryland, IEEE, pp. 751-754, 1993.

J.C. Principe, D. Xu, Q. Zhao, and J. W. Fisher III, **Learning From Examples with Information Theoretic Criteria**, University of Florida Gainesville, www.com/ Goggle/ Neural Network Learning rules, 2000.

P. E. Keller and A. F. Gmitro, **Design of Fixed Planer Holographic Interconnects for Optical Neural Networks**, Applied Optics Vol. 31, pp. 5517-5526, 10 September 1992.

P. E. Keller and A. F. Gmitro, **Operational Parameters of An Optoelectronic Neural Net Employing Fixed Planer Holographic Interconnections**, the World Congress of Neural Network, July 1993 (WCNN 93).

P. Moerland, E. Fiesler, and I. Saxena, **Incorporating LCLV Non- Linearities in Optical Multilayer Neural Networks**, Applied Optics VoL. 35, No. 26, pp. 1-10, September 1996.

W. H. Al-Buhrezi, "**Computer Implementation of Optoelectronic Artificial Neural Networks**", M.Sc Thesis, University of Technology, 2001.

Z. Chahramani, A. T. Korenberg and G. E. Henton, **Scaling in A Hierarchical Unsupervised Network**, Ninth International Conference on Artificial Neural Networks-University College London, 1999.

DESIGN AND IMPLEMENTATION OF APPLICATION PROGRAM MODULE FOR A POWER STATION SCADA SYSTEM

Assmaa A. Fahad
Lecturer
Baghdad University / College of Science / Computer Science Department

ABSTRACT

Supervisory Control And Data Acquisition (SCADA) is a commonly used industry term for computer based system allowing system operators to obtain real-time data related to the status of an electric power system over wide geographic area.the designed SCADA system consists of Five modules and it is designed to work with two layers: Client layer and Data server layer each in a separated PC. The APplication Program (APP) module is one of the SCADA modules that works as an interface between the I/O cards and the other SCADA modules. It collects the data from the I/O cards; through an I/O interface card; and passes it to the Process Data Interchange (PDI) module. And because of the APP module works in a separated PC, it communicate with PDI module using the PC serial port and High-level Data Link Control (HDLC) communication protocol. The designed SCADA modules are programmed as a multithread programs written using Visual C++ programming language. The APP module is built to consist of three threads: one for collecting data, another to perform the communication operation, and the third thread is responsible of controlling all APP operations. The overall SCADA system including the APP module is implemented to supervise the operation of a power station since 1999 and it proves a very fast response time and provides a good real-time reports.

الخلاصة

يعتبر نظام السيطرة الإشرافية وجمع البيانات (SCADA) من الأنظمة الشائعة الاستخدام للسيطرة على المعامل التي تعتمد الحاسبات الشخصية لتجميع المعلومات عن محطات الكهرباء بالوقت الحقيقي وعبر مسافات بعيدة. يتكون نظام SCADA ألذي تم بناءه من خمس وحدات رئيسية تعمل بطبقتين برمجية، طبقة الموكل وطبقة محضر البيانات وعلى حاسبتين منفصلتين. وحدة البرامج التطبيقية هي إحدى وحدات النظام الذي يعمل كواجهة بين حقول وحدات الإدخال والاخراج وباقي وحدات النظام. فهذه الوحدة مسؤولة عن عملية جمع البيانات من حقول وحدات الإدخال والاخراج وتمريرها إلى باقي الوحدات عن طريق وحدة تبادل المعلومات. ونظرا لعمل هذه الوحدة على حاسبة منفصلة فانها تستخدم وحدات الإدخال المتسلسل المتوفرة في الحاسبة كوسيلة اتصال مع وحدة تبادل المعلومات وباستخدام نظام التخاطب HDLC. تم برمجة وحدات نظام SCADA لتتكون من عدة خيوط برمجية باستخدام لغة

البرمجة Visual c++ . وحدة البرامج التطبيقية تم بناء برامجها لتتكون من ثلاث خيوط برمجية خصص أحدها لقراءة البيانات من وحدات الإدخال والإخراج فيما خصص الثاني للأشراف على عملية التخاطب وتبادل المعلومات من باقي وحدات النظام . أما الثالث وهو الرئيسي فقد برمج للسيطرة على عمل وحدات النظام ككل.

تم فحص عمل النظام بجميع وحداته للإشراف على عمل إحدى محطات الكهرباء منذ العام ١٩٩٩ وقد اثبت كفاءة عالية في الأداء من حيث وقت الاستجابة للإشارات الداخلة ونوعية التقارير التي تم تحريرها.

KEY WORDS:

APP module: APplication Program module
CRC: Cyclic Redundancy Check
FCS: Frame Check Sequence
HDLC: High-level Data Link Control
SCADA: Supervisory Control And Data Acquisition

INTRODUCTION

SCADA systems have made substantial progress over the recent years in terms of functionality, scalability, performance and openness such that they are an alternative to in house development even for very demanding and complex control systems as those of physics experiments.

SCADA system is not a full control system, but rather focuses on the supervisory level. As such, it is a purely software package that is positioned on top of hardware to which it is interfaced, in general via Programmable Logic Controllers (PLCs), or other commercial hardware modules. SCADA systems are used not only in industrial processes: such as steel making, power generation (conventional and nuclear) and distribution, chemistry, but also in some experimental facilities such as nuclear fusion. The size of such plants range from a few 1000 to several 10 thousands I/O channels. However, SCADA systems evolve rapidly and are now penetrating the market of plants with a number of I/O channels of several 100 K [Dennise 1987].

SCADA SYSTEM ARCHITECTURE

There are two basic layers in a SCADA system: the "client layer", which caters for the man machine interaction, and the "data server layer" which handles most of the process data control activities. The data servers communicate with devices in the field through process controllers. Process controllers, e.g. PLCs, are connected to the data server either directly or via networks or field buses [Dennise 1987] .

The data server polls the controller in an operation called scanning operation. Scanning operation is performed at a user defined polling rate which is differing for different parameters. The controller pass the requested parameters to the data server, time stamping of the process parameters is typically performed in the controller and this time-stamp is taken over by the data server.

Server-client communication is in general based on a publish-subscribe and event-driven using one of common communication protocols [Ghanim 1999]. The client application subscribes to a parameter which is owned by a server application and only changes to that parameter are then communicated to the client application.

The SCADA system software is multi-tasking software based upon a real-time database located in the server. Normally the SCADA system is built to consist of a number of modules each responsible of a specific function. The number of these modules and the functions they performed are depending on the requirements of the system the SCADA is built to control.

SCADA SYSTEM FUNCTIONS

The SCADA systems are designed to perform the following functions [Ghanim 1999]:

- **Access Control**
  Users are allocated to groups, which have defined read/write access privileges to the process parameters in the system and often also to specific product functionality.
- **Man Machine Interface (MMI)**
  The products support multiple screens, which can contain combinations of synoptic diagrams and text. They also support the concept of a "generic" graphical object with links to process variables. These objects can be "dragged and dropped" from a library and included into a synoptic diagram.
  Most of the SCADA products that were evaluated decompose the process in "atomic" parameters (e.g. a power supply current, its maximum value, its on/off status, etc.) to which a Tag-name is associated. The Tag-names used to link graphical objects to devices can be edited as required. The products include a library of standard graphical symbols, many of which would however not be applicable to the type of applications encountered in the experimental physics community.
  Standard windows editing facilities are provided: zooming, re-sizing, scrolling, etc. On-line configuration and customization of the MMI is possible for users with the appropriate privileges. Links can be created between display pages to navigate from one view to another.
- **Trending**
  The products all provide trending facilities (the parameters to be trended in a specific chart, number of trended parameters in each chart, etc). The trending feature is either provided as a separate module or as a graphical object (ActiveX), which can then be embedded into a synoptic display. XY and other statistical analysis plots are generally not provided.
- **Alarm Handling**
  Alarm handling is based on limit and status checking and performed in the data servers. More complicated expressions (using arithmetic or logical expressions) can be developed by creating derived parameters on which status or limit checking is then performed.
- **Logging/Archiving**
  The terms logging and archiving are often used to describe the same facility. However, logging can be thought of as medium-term storage of data on disk, whereas archiving is long-term storage of data either on disk or on another permanent storage medium. Logging is typically performed on a cyclic basis, i.e., once a certain file size, time period or number of points is reached the data is overwritten. Logging of data can be performed at a set frequency, or only initiated if the value changes or when a specific predefined event occurs. Logged data can be transferred to an archive once the log is full. The logged data is time-stamped and can be filtered when viewed by a user.
- **Report Generation**
  One can produce reports using SQL type queries to the archive, Real Time Data Base (RTDB) or logs. Although it is sometimes possible to embed EXCEL charts in the report, a "cut and paste" capability is in general not provided. Facilities exist to be able to automatically generate, print and archive reports.

.

THE DESIGNED POWER STATION SCADA SYSTEM

The System Architecture

The system is constructed around two PCs, the central computer and the I/O processor. The central computer PC is the client of the SCADA system, while the I/O processor PC is the data server. The I/O processor is connected to the process equipment by means of an I/O interface which provides interfacing between I/O cards bus line of the existing equipments and the bus of the I/O processor PC.

The designed system major parts are hardware I/O interface card and software modules that are designed to fulfill the system requirements. The software modules are:

- APplication Program module (APP).
- Process Data Interchange module (PDI).
- Data Base module (DB).
- Event Processing module (EP).
- Man-Machine Interface module (MMI).

Digital and analog data are acquired through the system hardware via the I/O interface driven by the I/O drivers in the APP module. The APP module residing in the I/O processor PC transmits the data to the PDI module in central PC via serial link using HDLC communication protocol.

PDI module manages the data exchange and transmission . After receiving the information, the PDI updates the data base files and transmits a transaction to the EP module. This occurs upon status change of digital data or if analog values exceed predefined limit, or in case of error occurrence in the system.

DB module manages the data base files and the logging and archiving process. The EP module is responsible of alarm handling and records the events to the operator.

MMI module display the information stored in the data base files either as diagrams or event lists or curves. MMI also provides initiation of user inquiries and commands, such as system diagnostic and test requests. Event lists, reports and other messages can be directed to a printer if required.

The scope of this paper is to design APP module in SCADA system.

The Designed APP Module

As mentioned before the APP module is responsible of providing the plant's status to another modules. Therefore it is structured to perform the following functions:

- Provide I/O driver routines for I/O interface card.
- Perform scanning operation .
- Interrupts handling.
- Transmit field status to central PC.
- Perform Diagnostic operation.

I/O Driver Routines

Driver routines for I/O interface card are required as a major part of any APP module. To achieve highest accuracy the structure of these drivers should be such fast response time and minimum software overhead. Therefore these programs should be simple, flexible and with minimum code.

To write I/O driver routines the structure of the I/O field and the data types this I/O field provides must be specified.

The designed SCADA system is implemented in a hardware with six thermal units connected to the computers. Inputs from the plant are interfaced directly to seven separated subsystems for analog inputs and digital inputs collection. One subsystem is foreseen for each of the six power units, and one dedicated to common services data collection[Baiji Manual 1989].

Several I/O driver routines are written to read different types of data. The system accepts plant data in form of analog inputs, digital inputs or pulse inputs and generates output digital signals to perform loop-back test to the I/O cards.

For analog inputs the plant includes 14 boards of 16 channels each. These inputs are applied to an Analog to Digital Converter (ADC) via multiplexers. At the end of conversion the

ADC generates an interrupt signal after which APP module can read the analog data from ADC buffer.

The digital inputs arranged in 32 boards of 16 channels each. These channels are known as Digital-scan channels. Two driver routines are written to read data from these channels, one to read all 32 boards at one time and another to read single specified board.

Pulse input channels are digital input with interrupt, changing the status of any channel will cause an interrupt signal; these channels are known as Digital-interrupt channels. The signals connected to these channels are with the highest priority in the system. In this type the inputs arranged in 16 boards of 12 channels each. The data from these channels are reads either when an interrupt occurs from these channels or by using an I/O driver routine written to allow selective reading of these channels at normal time (no interrupt).

Scanning Operation

In this operation the APP module reads the status of the plant from the I/O field. At the initialization of SCADA system all channels are scanned in order to record the initial state of the plant. After that the scanning operation is performed periodically depending on the rate of change of data on these channels. For the designed system and depending on this rate of change the analog channels are divided into seven types: one second, two seconds, five seconds, ten seconds, twenty seconds, thirty seconds, and sixty seconds data change rate channels. In accordance with this classification the APP module perform the scanning operation, some channels scanned every one second another scanned every two seconds and so on. Digital–scan input channels are all with the same data rate change, every one second, so they are all scanned every one second. Digital-interrupt channels are not scanned for there nature of operation. All channels are scanned, but only the changed one will transmitted to central PC.

During the scanning operation, the APP module reads the status of the channel and compare it with the old one stored in the DB, if there is a change the channel address, the new status of the channel, and the time stamp (if need) are stored in a circular queue in order to be arranged in a frame and send to central PC.

Three circular queues are defined to store the status of the channels one for each type (Digital-scan, Digital-interrupt, and Analog). Because of APP module is built as a multithread program these queues may accessed from more than one thread at the same time. To synchronize the accessing of these queues without any error, these queues are defined as a critical section area [Silberschatz 1998].

With Digital-scan channels and analog channels there is no needs for a time stamp because they scanned periodically, therefore the status of each channel is represented in only three bytes (one for channel address and two bytes for channel status in its digital form). With digital-interrupt channels it is very important to fix the time of the event, therefore the status of each channel will represented in five bytes (6-bit channel address, 4-bit modified point in this channel, 1-bit for point status, 1-bit not used, and 28-bit for time stamp which is fixed in millisecond), Fig.(1).

0		7
Channel Address (6-bit)	Point status (1-bit)	not used (1-bit)
Modified Digital Point No. (4-bit)	Millisecond (4-bit)	
Millisecond (8-bit)		
Millisecond (8-bit)		
Millisecond (8-bit)		

Fig (1) Digital-interrupt information record

Interrupt Handling

One of the most important and critical operation needs to be handled by APP module is handling the interrupt signals generated by the hardware components. The designed APP module handles four interrupt signals:

- Interrupt signal from A/D converter: When A/D completes his operation it sends an interrupt signal, the Interrupt Service Routine (ISR) handles this interrupt signal will read the data from A/D buffer and initiate, if any, next analog scanning.
- Interrupt signal from system timer: The APP module programs Intel 8254 system timer to generate an interrupt signal every 150ms. Accordingly the APP will update the time and the date of the system. The APP module use this time to synchronize the communication operation with central PC, fix the time of scanning operation, record the time of the events, and reset the watch dog hardware timer. The watch dog timer is a part of the I/O interface card that used to insure continuous operation of the system. This timer must resets every 200ms, if for some reason the system is stopped and the watch dog timer not resets, this timer will cutoff the power supply of the I/O boards.
- Interrupt signal from any changed Digital-interrupt channel: The ISR handles this interrupt signal will read the new status of this channel and records it, with all other information, in a digital-interrupt circular queue.
- Interrupt signal from serial port: PC serial port is used to communicate the data with central PC. It is programmed to send an interrupt signal when the receiver buffer is full. The ISR reads the characters in the buffer and return the control to APP main routine.

Transmit Field Status to Central PC

This part of APP module is designed to transmit the status of the channels from APP module to central PC through PDI module. The two PCs, I/O processor PC and central PC are placed in two different places and they are connected through RS232 serial port. HDLC communication protocol is used to control the communication operation between these PCs [Halsall 1996].

Different types of messages are defined to handle the communication operation some of these messages are used to control the communication operation and another are used to hold data. In HDLC protocol both data and control messages are carried in a standard frame format, Fig (2) [Halsall 1996].

Start of frame delimiter (8-bit)	Frame Header (16 bit) 8-bit	8-bit	Information Field 0.. N	Frame Check Seq. (16-bit)	End of frame delimiter (8-bit)
Flag	Address	Control	Information	FCS	Flag

Fig (2) Standard HDLC frame format

The designed SCADA system uses the Flag byte to define the start and the end of the message. The address field is used to check for data transparency (the appearance of flag sequence in frame contents) because the connection between the two PCs is point-to-point [Halsall 1996].

The Control field is used to define the frame type: Unnumbered frame, Information frame, and Supervisory frame. For each type the designed SCADA system use the control field as in Fig (3).

	1	2	3	4	5	6	7	8
Information frame	0	N(S)			P/F	N(R)		

Where N(S) : send sequence number
N(R) : receive sequence number
P/F : Poll / Final bit

	1	2	3	4	5	6	7	8
Supervisory frame	1	0	S		P/F	N(R)		

Where S : Receiver Ready – RR
Receiver Not Ready - RNR
Reject – REJ
Selective Reject – SREJ

	1	2	3	4	5	6	7	8
Unnumbered frame	1	1	M		P/F	M		

Where M : <u>Unnumbered commands</u>
Set Normal Response Mode (SNRM)
Frame Reject (FRMR)
Disconnect (DISC)
<u>Unnumbered responses</u>
Unnumbered Acknowledge (UA)
Frame Reject (FRMR)
Disconnected Mode (DM)

Fig (3) SCADA system HDLC Control Field formats

Frame Check Sequence (FCS) field is a 16-bit Cyclic Redundancy Check (CRC) computed for the complete frame contents enclosed between the two flag delimiters. This field is computed and transmitted with each message in order to detect any error may occur during the communication operation. The CRC error detection method is used to increase the reliability of the system [Halsall 1996].

These information messages and control messages are implemented in both APP and PDI modules. The APP module receives a number of control messages from PDI module that control the communication operation between the two PCs. Five link modes are defined in communication operation: Link-establishment mode, Time-sync mode, Data-request mode, Diagnostic mode, and Link-disconnected mode.

The data communication operation is activated when the link is established upon request from central PC. When the link is established between the two PCs a time-sync mode is activated in which central PC sends a time-sync message to synchronize the time in two PCs. When APP receives a time-sync message it updates the timer counters in I/O processor PC accordingly. Time-sync mode is activated from time to time to always insure the synchronization of the time between the two PCs.

Data-request mode is activated after time-sync mode, in this mode APP module begins a transmission of data messages. One information message can hold the status of more than one channel but with the same type, for example one message can hold the status of all changed Digital-scan input channels that scanned at specific time. With any transmission error, APP module retransmits the last message for three times and if there is still an error a Link-establishment mode is reactivated to check the status of the connection between the two PCs. With any error occurs during the communication or scanning operation the APP module also sends an error message such as: Dig-scan-buffer-full, Address-error-scan-seq., Analog-buffer-full, etc to central PC in order to recover it, Fig (4).

8-bit	8-bit	Control				1-bit	10-bit	5-bit	16-bit	8-bit
Flag	Address	0	N(S)	P/F	N(R)	Error	Channel Address	Error code	CRC	Flag

Fig (4) SCADA system error message format

After sending any error message the APP module continues with his normal operation. The PDI module passes this error message to EP module in order to display a suitable report to the operator describing the problem in the system.

Diagnostic Operation

Checking the performance of the I/O boards in the I/O field is called diagnostic operation. Diagnostic operation is performed by using special diagnostic cards. These cards are used to route a defined pattern of zeros and ones to the I/O boards. The APP then reads the data from the I/O boards and compares it with the written one. The diagnostic operation is activated by central PC when there are a number of error messages (such as Address-error-scan-seq.) are received from APP module.

When APP module receives a diagnostic request message from central PC it suspends all the operations and changes the mode of operation to diagnostic mode and begins the diagnostic operation. After completing the diagnostic operation the APP module sends the diagnostic results in messages to central PC and resumes his operations. The EP module in central PC will display a report about the diagnostic result so the operator can make a decision about the system accordingly.

DISCUSSION AND CONCLUSION

In spite of the complexity of the SCADA system and very large and different types of signals received from the I/O fields, the designed SCADA system improves a good performance when it is implemented to supervise the operation of a power station since 1999. The system

performance measure is by examining the CPU usage and the memory usage. During the operation of the designed SCADA system the PCs are work with minimum CPU and memory usage.

Programming SCADA modules as a multithread system simplify running these modules concurrently and simplify the operation of writing very efficient programs with maximum use of the CPU.

HDLC communication protocol is very suitable communication protocol for such SCADA system for his flexible frame format that makes the system provides different types of control and information messages and a different number of operation modes that help the programmer to cover all the requirements of the system.

REFERENCES

"Data logger and alarm system for Baiji power station", Technical Manual, 1989.

Dennise J. G. and Henry T. D. " Supervisory Control and Data Acquisition", proceeding of the IEEE, Vol.7, No. 12, December 1987.

Ghanim, Z. N. " Design and Implementation of Small SCADA System", a master theses, Baghdad University, College of Engineering, 1999.

Halsall, F. " Data Communications, Computer Networks and Open System", fourth edition, Addison Wesley, 1996.

Silberschatz A. and Galvin P. B. "Operating system concepts", Fifth edition, Addison-Wesley, 1998.

ESTIMATION OF CRITICAL BED DEPTH IN FIXED BED OF GRANULAR ACTIVATED CARBON

Prof. Dr. Abbas H. Sulaymon
Baghdad University
Eng. College

Waleed M. Abood
Head of Engineers
Energy and Environment
Research Center

ABSTRACT

The aim of this study is estimating the critical bed depth in adsorption process through a fixed-bed of granular activated carbon at different bed depths of 0.03, 0.05, 0.08and 0.11m at influent furfural concentration in waste water of 0.2 kg/m^3, with constant flow rate of (16.66) × 10^{-5} m^3/min and adsorbent particle size (0.5-1.5) mm. the changing of flow rate and furfural influent concentration had been studied to determine their effects on the critical bed depth value by using bed depth- service time method (BDST). Length of unused bed (LUB) and length of equivalent section of bed had been estimated mathematically during process of the adsorption at different bed depths and during changing the flow rate(8.3× 10^{-5} m^3/min) and influent concentration at same bed depth (0.05m).

الخلاصة:

الهدف من هذه الدراسة هو احتساب عمق الحشوة الحرج لعملية الامتزاز في حشوة ثابتة من الفحم المنشط الحبيبي ذات اعماق متنوعة هي (٠,٠٣، ٠,٠٥ ، ٠,٠٨، ٠,١١) م بثبوت كل من معدل الجريان بقيمة (١٦,٦٦) × 10^{-5} م^3/ دقيقة ، وتركيز الفورفورال الداخل في المياه المختلفة (٠,٢) كغم/م^3 وحجم حبيبات الفحم (١,٥ – ٠,٥) ملم. تغير معدل الجريان وتغير تركيز المادة الداخلة تم دراستها لغرض تحديد تأثير عمق الحشوة الحرج باستخدام طريقة عمق الحشوة – زمن الخدمة) . طول الحشوة الغير مستخدم والطول المكافئ للحشوة المستخدمة تم تحديدها رياضياً خلال عملية الامتزاز للظروف اعلاه ودراسة تأثير تغير معدل الجريان وتركيز المادة الداخلة عند ثبوت اعماق الحشوات المستخدمة.

KEY WORDS: Adsorption, furfural, break through, exhausting point, length of unused bed, length of equivalent section, critical bed depth.

INTRODUCTION

In fixed – bed, dynamic adsorption system a liquid rich in sorbable component (adsorbate) flows through bed or zone containing adsorbent particles (Lukchis, 1973). The contact time between adsorbate and adsorbent is limited by the rate and geometry of the bed (Crittenden and Thomas, 1998). Granular activated carbon (GAC) in fixed – bed column is generally preferred to use in powdered form, where continuous application is required. GAC allows a more complete use of adsorption capacity of the carbon and prevent clogging in the bed thus reducing make-up costs (Casey, 1992).

Empirical or short – cut methods were still used extensively for the design of fixed bed column which are, length of unused bed (LUB), mass transfer zone length (MTZL), empty bed contact time (EBCT), bed depth – service time (BDST) and capacity of break point (Crittenden and Thomas, 1998). Bed Depth Service Time (BDST) method is used extensively by the water industry and can be applied to other situation where (BDST) is a common method for evaluating pilots column data from graphical analysis of the break through curve and analysis of granular – carbon system (Richard, 1998) (Ray, 1973). Length of unused bed (LUB) represents the part of the total bed at the end equilibrium zone (break through time) therefore increasing the bed size beyond the length of the mass transfer zone (MTZ) (critical bed depth) (Crittenden and Thomas, 1998) able equilibrium capacity – adsorbent that provides 100% use due to increasing break through time (Lukchis, 1973)that help to determine the active bed.

EXPERMINTAL WORK

Down flow fixed – bed vertical polyvinyl chloride (PVC) column with inside diameter (0.04) m was used for the experiments at different bed depths of 0.03, 0.05, 0.08and 0.11 m and constant flow rate (16.66)$\times 10^{-5}$ m^3/min , influent furfural concentration 0.2 kg/m^3, carbon particle size (0.5–1.5) mm. Adsorbate is a furfural (OSHA, 2000) in industrial waste water obtained from Al- Dora refinery treatment plant after the primary treatment stage (Akrawi, 1988). Adsorbate is granular activated carbon (Provided by UniCarbo CO. Italia) with bulk density 460-480 kg/ m^3 and mesh number 0.4–1.6 mm with high surface area 1100 – 1200 m^2/gm (Carbo- chem`s report, 2004) (Lyderson, 1983). The samples were taken from the bottom of the column at each 15 min. each sample (5 ml) was analyzed by using calorimetric (type Jenway, 6030, UK.) at wave length 430 nm after adding 65 ml of 27.5 vol.% ethanol, 5 ml of 10 vol. % .acetic acid and 0.5 ml of distilled aniline as an indicator (Foste and Leslie, 1971).

RESULT AND DISCUSSION

Break through curves:

Break through curves were plotted between effluent furfural concentration C/Co vs. time at different bed depth and constant other variables **Figure (1)** (Abood, 2005)

Fig (1) shows the increasing the bed depth leads to the increasing break through time (Tb) with clear S – shape of curve for mass transfer that is due to the provide of extra surface area for more adsorbate to carry on. (Martin and Al- Bahrani, 1978).

- Bed Depth Service Times:

Horizontal lines of (C / Co = 0.1) and (C / Co = 0.9) were plotted as shown in **Figure (1)** in order to determine of break through point and exhaustion point respectively for each curve **Table -1-**.

Table (1) Bed Depth Service Time

Bed depth (m)	0.03	0.05	0.08	0.11
Break point time (min)	15	120	277.5	435
Exhaustion point time (min)	187.5	285	457.5	625

Fig(2) was plotted as service time vs. bed depth at flow rate $(16.66\times10^{-5})m^3/min$, influent concentration 0.2 kg/m^3 and particle size (0.5–1.5) mm.

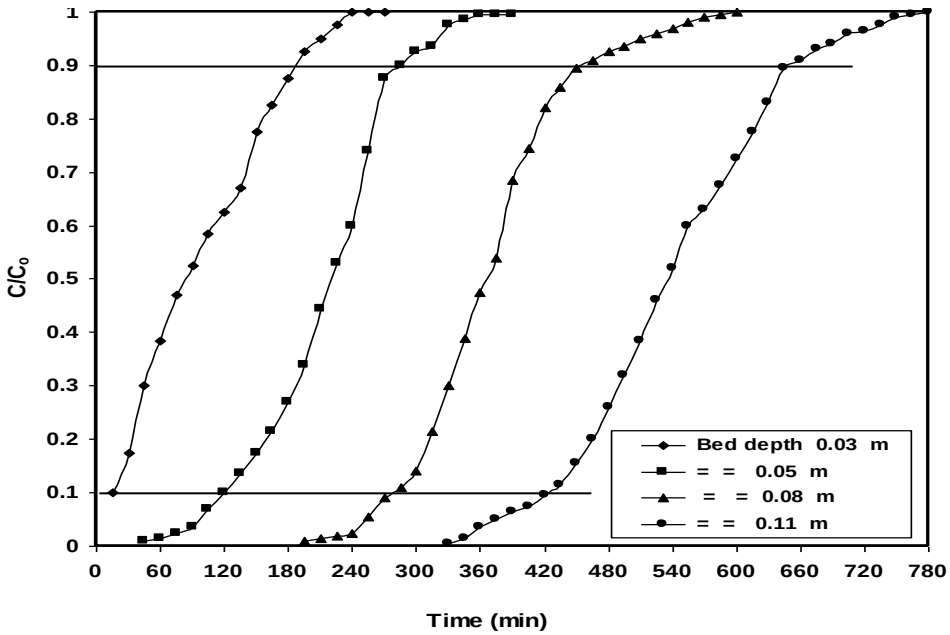


Fig. (1) Plot of (C/C_o) vs. time at different bed depth

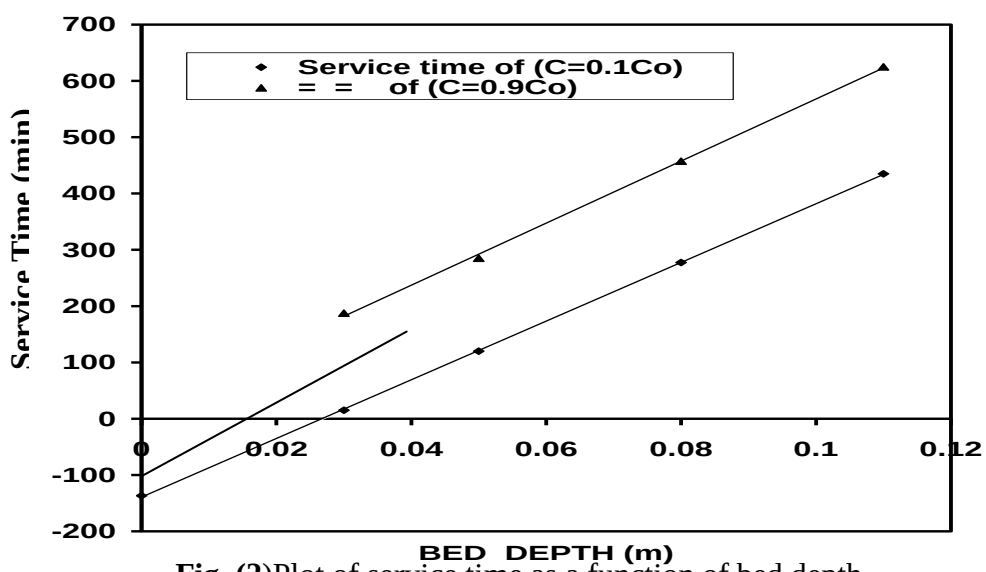


Fig. (2) Plot of service time as a function of bed depth

By fitting the lines which represent the best result in **Fig (2)** the regression equation was determined as follow:

$$\text{Time} = A X + B \quad \ldots\ldots\ldots\ldots(1)$$

Where: X= Bed depth m.
A = the slope (min/ m) or (day / m) = 4 day / m
B = intercept of (C / Co = 0.1) line with Y – axis (Service time) (min) or (day) = - 0.095 day.
The intercept of (C / Co = 0.1) line with the X – axis (bed depth) represents the critical bed depth which approximately with the lowest value of experimental bed depth which means the mass transfer zone length without equilibrium zone (Crittenden and Thomas, 1998) (Richard, 1998)

EFFECT OF CHANGING FLOW RATE ON CRITICAL BED DEPTH

For changing the operating criteria like flow rate at constant influent concentration, particle size and different bed depth, equation (1) becomes as follow:

$$\text{Time} = A' X + B \quad \ldots\ldots\ldots\ldots(2)$$

At same B.

$$A' = A (Q_{old} / Q_{new}) \quad \ldots\ldots\ldots\ldots(3)$$

A′ = new slope of (C / Co = 0.1) line (day / m) at flow rate (8.3×10^{-5})m^3 / min the A′ value equal to 8 day / m at constant B value. **Fig (3)**.
Q = flow rate (m^3/min).

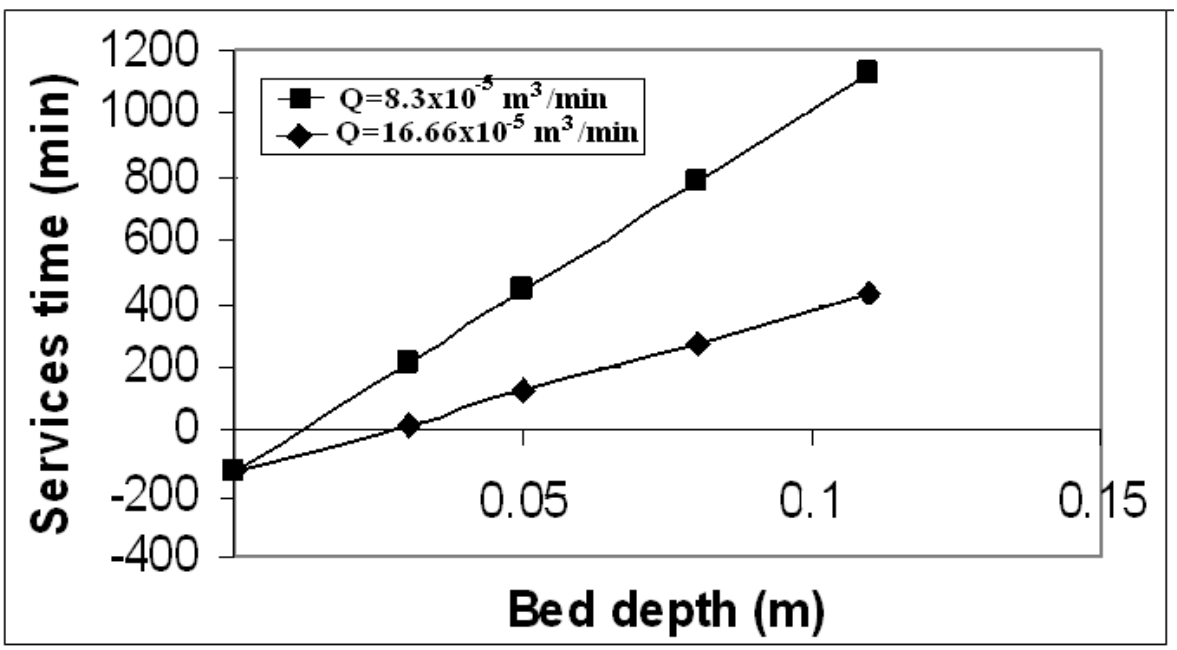


Fig (3) Bed depth service time at different flow rate

(C/Co = 0.1) line in **Fig (3)** which shows the intercept of the line with x– axis to determine the new critical bed depth, decreasing flow rate will decrease the critical bed depth this is due to the more

contact time at constant cross section area and bed depth and increasing break through point time therefore the flow rate per section area must be not more than (0.132 m^3/m^2.min (Cassy, 1992)**Table – 2 –**.

Table – 2 – Critical bed depth at different flow rate

Flow rate 10^{-5} (m^3/ min)	**A (day / m)**	**B (day)**	**Critical bed depth (m)**
16.66	4	-0.0 95	0.024
8.3	8	- 0 .095	0.012

Effect of Influent Concentration on the Critical Bed Depth

Changing the influent concentration at different bed depths (0.03, 0.05, 0.08, 0.11) m ,constant flow rate (16.66) × 10^{-5} m^2/min and particle size (0.5 – 1.5)mm, the regression equation (1) will become as follow :

$$\text{Time} = A' X + B' \quad \ldots\ldots\ldots\ldots(2)$$

Where:
A´ = new slope = A (Co /C′o)
new intercept with Y – axis B′ = B (Co /C′o)[ln (Co/Ce -1)/ ln (C′o/ Ce -1)]
Where :
Co = old concentration (kg / m^3)
Co′= new concentration (kg / m^3)
Ce = effluent concentration equal to 0.9 influent concentration
Figure (4) shows the new regression equations and old one where **Table -3-** shows A′ , B′ and critical bed depth for each influent concentration at different bed depths.

Table -3- Critical bed depth values

Co (kg/ m^2)	**A (day/m)**	**B (day)**	**Critical bed depth (m)**
0.3	2.667	-0.0633	0.024
0.2	4	-0.095	0.024
0.1	8	-0.19	0.024
0.05	16	-0.38	0.024

Table -3- shows that the critical bed depth values are same as shown in **Figure (4)** there are same intercepts lines (C/Co= 0.1) with the X– axis.

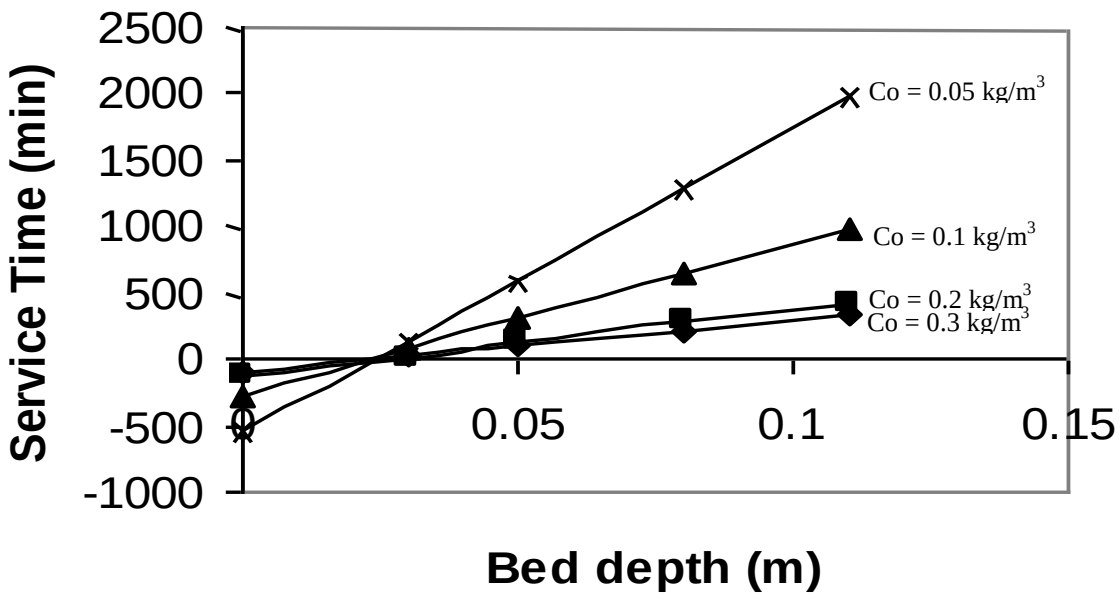


Fig (4) Bed depth Service Time at Different influent concentration

Length of Unused Bed (LUB):

Determination of length of unused bed (LUB) depend on break through times Tb at (C/Co = 0.1) and stiochemitric times Ts where Ts consider the time of mass transfer zone front, equation (5) had been used to estimated (LUB) value:

$$LUB = L_o (1 - Tb / Ts) \quad \ldots\ldots\ldots\ldots(5)$$

$$Ts = Tb + (Te - Tb) / 2 \quad \ldots\ldots\ldots\ldots(6)$$

$$Time = A' X + B' \quad \ldots\ldots\ldots\ldots(2)$$

$$Time = A' X + B' \quad \ldots\ldots\ldots\ldots(2)$$

Where :

Lo = length of original bed depth (m)

Te = exhaustion time at (C/C_o = 0.9) (min)

The length of equivalent section (LES) at break through time represent the length of bed when the process of adsorption must be stopped for replacing the bed or regeneration and its calculated as follow (Likchis, 1973).

$$LES = L_o - LUB \quad \ldots\ldots\ldots\ldots(7)$$

$$Time = A' X + B' \quad \ldots\ldots\ldots\ldots(2)$$

Tabl l depth when (0.03) m approximately equal to the value of critical bed depth as shown in **Figure (2)**, therefore the bed depth must be at least (1.5-3) times of the mass transfer zone length (MTZL) (Lukchis, 1973). **Figure (5)** shows the increasing of percentage used bed with increasing bed depth

Table -4- shows the values of LUB and LES at different bed depth and constant other operation condition.

Bed depth (m)	Tb (min)	Ts (min)	LUB (m)	LES (m)	% used bed
0.03	15	101	0.255	0.0045	14
0.05	120	202	0.2	0.03	60
0.08	277.5	367.5	0.2	0.06	75.5
0.11	435	530	0.2	0.08	82

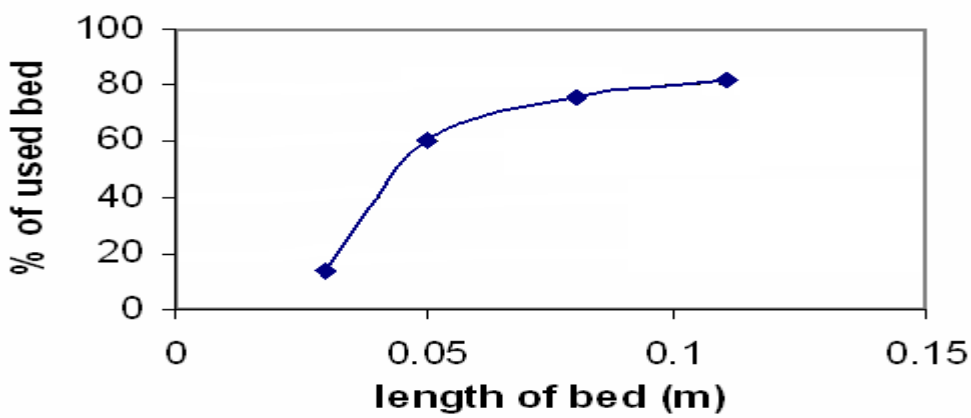


Fig (5) Percentage of used bed

At same bed depths when flow rate had been changed from (16.66 to 8.3)×10^{-5} m^3/min theoretically the LUB and LES will change to as shown in **Table -5-** at (0.05) m bed depth, (0.2)kg/m^3 concentration and (0.5-1.5) particle size.

Table -5- LUB & LES values

Flow rate m3/min	Bed depth (m)	Ts (min)	Tb (min)	LUB (m)	LES (m)
16.66 × 10^{-5}	0.05	120	202	0.02	0.03
8.3 × 10^{-5}	0.05	360	477.5	0.012	0.0376

The value of LUB is approximately equal to value of intercept (C / Co =0.1) line in **Figure (3)** at flow rate (8.3 × 10^{-5}) m^3/min which approve that decreasing flow rate increasing break through time.
There is no effect of changing influent concentration on the values of LUB and LES these due to the constant time between the adsorbate- adsorbent is limiting by the rate and geometry of bed (Lukchis, 1973) where **Figure (4)** shows no effect of influent concentration or the critical bed value.

CONCLUSIONS

- Break through time increase with the increase of bed depth and the decrease of flow rate and influent concentration
- The given results from bed depth – service time method approximate with experimental results of the length of critical bed depth.
- Length of equivalent section (LES) increases with the increase of the bed depth and the decrease of the flow rate.
- There is no effect of changing the influent concentration on the critical bed depth at constant flow rate in bed depth service time methods analysis.
- The values of length of unused bed (LUB) is constant even though adds more size of bed was added beyond mass transfer zone and decreases when flow rate decreases.

REFERENCES

- Abood W. M., 2005, M. Sc. Thesis, university of Baghdad
- Akrawi R. F., 1988, M. Sc. Thesis, university of Baghdad.
- Carbochems report, 2004, " activated carbon", at <www. Carbochem.com/activatedcarbon101.html.>
- Casey I. J., 1992, "Unit Treatment Processes", Wiley services, New York.
- Crittenden B. and Thomas J. W., 1998, "Adsorption Technology and Design", Butter Worth, UK.
- Foste R. and Lesile W., 1971, "Encyclopedia of Industrial Chemical Aanalysis", John Wiley, New York, vol. 13, p. 232-239.
- Lukchis A. L., 1973, J. Chem. Eng. Science, June, p. 111-116.
- Lyderson A. L., 1983, "Mass Transfer in Engineering Particle", John Wiley & Sons New York.
- Martin R. J. and Al- Bahraini. S., 1973, Water Research, vol.12, p.879-888.
- Ray A. A., 1973, J. Chem. Eng., August 20.
- Watts R., 1998, "Hazardous Wastes", John Wiley, New York.

NOMENCLATURE

A Slope of regression equation (day/ m)
B Intercept of regression equation with y- axis (day)
C Effluent concentration at given time (kg/m^3)
Ce Equilibrium concentration (kg/m^3)
Co Influent concentration (kg/m^3)
L Length of carbon bed (m)
Q Flow rate (m^3/ min)
Tb Break through time (min)
Te Exhaustion time (min)
Ts Sstoichiometric time (min)

TREATMENT OF DEPLETED URANIUM CONTAMINATION IN SOIL BY USING SODIUM BICARBONATE SOLUTION

Ahmed A. Mohammed, Ali Sh. M. H. Al-Attar

Environmental Engineering Department, College of Engineering, University of Baghdad.

ABSTRACT

The Depleted Uranium contamination in soil was treated with chemical leaching method by using sodium bicarbonate with respect to the effect of several variables (Time, Temperature, Bicarbonate Concentration, Carbonate/Bicarbonate Ratio, Oxidative Reagent Effect, pH, Soil/Solution Ratio and Rinsing Effect after treatment) in order to decontaminate or remove Depleted Uranium to acceptable regulatory levels.

The objective is to reach a selectively extracted Depleted Uranium by using a soil washing/extraction without generating a secondary waste which would be difficult to manage and/or dispose off. Results of Depleted Uranium removal efficiency were ranged from (35.4-88.25) %.

الخلاصة

التلوث الناجم عن اليورانيوم المنضب في التربة عولج بإستخدام طريقة الغسل الكيمياوي بواسطة بيكاربونات الصوديوم أخذاً بنظر الإعتبار تأثير متغيرات عديدة (الزمن، درجة الحرارة، تركيز البيكاربونات، نسبة الكاربونات الى البيكاربونات، تأثير العامل المؤكسد، الأس الهيدروجيني(pH)، نسبة التربة الى المحلول، تأثير الشطف بعد المعالجة) لازالة او معالجة التلوث الناجم عن اليورانيوم المنضب وصولا الى نتيجة تكون ضمن الحدود المسموحة.

الهدف كان ايضا انتقاء طرقة معالجة بالغسل الكيمياوي لا تولد تلوثا ثانويا كبيرا يصعب التخلص منه.نتائج كفاءة الأزالة تراوحت ما بين (٣٥،٤ الى ٨٨،٢٥)%. بالإضافة الى دراسة تأثير ثلاثة محاليل(حامض الكبريتيك، حامض الستريك وبيكاربونات يتبعها حامض الستريك) في عملية المعالجة، وكذلك دراسة تاثير المعالجة بالإذابة على(الحديد و الرصاص) في التربة.أخيراً، تم إيجاد علاقات رياضية خاصة بظروف التجارب التي اجريت ما بين تركيز اليورانيوم المنضب والمتغيرات المذكورة أعلاه.

INTRODUCTION

When measuring isotopic ratios in environmental samples it is important to realize that uranium may sometimes become depleted (or enriched) in some of its isotopes due to natural processes such as chemical weathering. Depleted uranium (DU) is a by-product from the process used to enrich natural uranium ore for use as fuel in nuclear reactors and nuclear weapons.

In weapon use ,when penetrator impact on ground surface ,a portion of its DU mass is transformed into aerosols or fine particles and thrown into the surrounding air .These aerosols and fine particles are normally depleted in measurable quantities on the surroundings ground or on other surfaces within about 100m from impact[1].

It is important to solve DU contamination problem in soil. There are several methods of treatment, basically are classified as physical and chemical, soil washing in a conventional sense is based on a physical separation such as screening classification (separation of soil particulate according to their settling velocities). Chemical extraction processes characteristically used to remove uranium from uranium ores are either acid or carbonate based extractions.

For acid extraction, sulfuric acid, which is less expensive than nitric, is the most commonly used acid. Other chemical extractions like carbonate extractions are highly selective for uranium.

The efficiency of the carbonate extractions is based on the formation of sodium or ammonium uranyl tri- and di-carbonate [UO_2 $(CO_3)_2$ and UO_2 $(CO_3)_3$], highly stable, water-stable and anionic complexes. Oxidants such as potassium permanganates may be used to increase the extraction efficiency. [2]

In the present work of treatment of DU contamination various variables were examined in leaching experiments such as time, temperature, soil to solution ratio, pH, $NaHCO_3$ concentration, carbonate/bicarbonate ratio, oxidizing reagent and rinsing effect, finally best conditions were obtained to give the best treatment.

- **ALKALINE LEACHING OF URANIUM**

Alkaline leaching is used only for materials, such as carbonates, which would consume a wasteful amount of acid. Its main advantage is the relatively non-corrosive properties of the solutions employed and the fact that few impurities are dissolved along with the uranium [3].

From these solutions:

1. Carbonate leaching involves the formation of various highly soluble U-carbonate complexes which are not likely to absorb to negatively charged soil constituents [4]. This allows high concentration of uranium in the leachate solution [5].

Some of the key reactions in this process are:

$$CO_3^{-2} + UO_2^{+2} \rightarrow UO_2CO_3 \quad (1)$$

$$2CO_3^{-2} + UO_2^{+2} \rightarrow UO_2(CO_3)_2^{-2} \quad (2)$$

$$3CO_3^{-2} + UO_2^{+2} \rightarrow UO_2(CO_3)_3^{-4} \quad (3)$$

$$6CO_3^{-2} + 3UO_2^{+2} \rightarrow (UO_2)_3(CO_3)_6^{-6} \quad (4)\ [6]$$

(Mason et al., 1997)[7] Used 0.5M sodium bicarbonate as the dominant reagent, was able to achieve uranium removal of (75-90%).

(Duff. et al., 1997)[8] Found that 0.5M sodium bicarbonate with oxidative compound indicated the overall efficiency of removal (52%) for different soil samples.

(Francis et al., 1994)[9] in their leaching design, used carbonate leaching media showed that >90% of the uranium can be removed from their soil samples.

(Mattus et al., 1993)[10] Found that carbonate extractions generally removed (40-90%) of the uranium from different soil samples.

(Timpson et al., 1994)[11] used the ultrasound treatments combined with carbonate extractions; the result was much as 90% of uranium removed.

(Francis et al., 1997)[2]Utilized carbonate/bicarbonate solution in the treatment, and this solution gave them good results in the removal of uranium.

(Elles et al., 1993)[12] also used the carbonate/bicarbonate solution in the removal of uranium from contaminated soil.

EXPERIMENTAL WORK

- **Soil Preparation**

The soil samples were dried by exposing to air for four days [13] with respect to good save conditions. After drying soil samples, the impurities (like small metal pieces and plants) were removed and then crushed and sieved (2mm Diameter) [14,15] before the treatment ,the samples were left for (28-30) day to reach the equilibrium state for the radionuclides that exist in soil [13,16].

- **Equipment and Materials**

1. Beakers & Plastic cups.
2. Filtration Paper. (Whatman 41).
3. Distilled Water.
4. Sensitive Balance (Mettler AC 100).
5. pH Meter (Expandable Ion Analyzer EA B40).
6. Thermometer.
7. Water Bath (Labsco, Germany).
8. Mechanical Compressing Device (Wabash, made in USA)

- **Solutions Preparation**

Solutions that used for uranium leaching process were prepared as below:

Bicarbonate solution

The bicarbonate solution was prepared with different range of molarities (0.2, 0.3, 0.5, 0.7 and 0.8). The weight of bicarbonate powder was taken by using the sensitive balance, and calculated by using the following equation (for 1liter of distilled water):

M = weight of compound (w)/molecular weight of compound (M.W)

The results were tabulated in **Table (1).**

Table (1):$NaHCO_3$ preparation and pH value.

Weight (gram)	Concentration (molar)	pH
16.8	0.2	8.2
25.2	0.3	8.1
42	0.5	8.06
58.8	0.7	8.02
67.2	0.8	8.03

Each of these weights were diluted with 1 liter of distilled water, then were placed in bottles and labeled with solution name, preparation temperature and date of preparation. The other solutions were prepared in the same manner by using the molarity & normality equations for preparation.(for 1 liter of distilled water)

Normality = Weight of compound (W)/equivalent weight of compound/volume

eq.w= Molecular weight (M.W)/n

Where n = number of protons (in acid-base reaction), or total change in oxidation number of compound (in redox).

Table (2) below shows the solutions, concentrations and pH

Table (2): Solutions preparation

Solutions	Concentrations	pH
Carbonate/Bicarbonate	37 g carbonate+18.5 g bicarbonate/liter	9.03
Carbonate/Bicarbonate	18.5 g carbonate+37g bicarbonate/liter	9.03
Carbonate/Bicarbonate	18.5 g carbonate+18.5 g bicarbonate/liter	8.45

- **Leaching Process**

The three samples were detected and the highest concentration was used in leaching experiments (detection process will be explained later).

All experiments were prepared in the same manner, where 0.8 g of soil were added to batch leaching solution then the solution-soil combination were left for a period. The

soil samples solution were filtered by filtration paper, then the samples were dried, prepared to detection procedure and labeled with sample name and other experiment conditions.

A. Time Set

S1 were used in all experiments, because of its high concentration compared with the other samples. This set consisted of five samples, each sample was represented a period of leaching.

The five experiments in this set were prepared as follows:

Soil sample were placed in the plastic cup, then 0.5M of $NaHCO_3$ were added with 1:20 soil to solution weight ratio. This batch solution was left for 2,3,4,5 and 7 days. (at temperature = 15 °C).

B. Temperature Set

Soil sample were placed in plastic cup with 0.5M $NaHCO_3$, soil to solution weight ratio was 1:20, the sample left in flask in water bath for 3 hours and half, this procedure were repeated five times, each one represented a temperature degree (20, 30, 40, 50 and 60 °C).

C. Concentration Set

In this set five molar values have been taken (0.2, 0.3, 0.5, 0.7 and 0.8).Soil samples were used with each of the five solutions for 4 days, at 23C^o , soil to solution weight ratio 1:20, then the samples were filtered, dried and labeled as the previous sets.

D. Carbonate/Bicarbonate Weight Ratio Set

The experiments are the same, except the experiment conditions were at temperature= 26 °C)., soil to solution ratio = 1:30, the samples were labeled with respect to the solutions' preparation (with Carbonate/Bicarbonate Ratio =1:1, 1:2 and 2:1).

E. pH Set

A 0.5M $NaHCO_3$ was placed in 4 plastic cups, soil samples were added to each one. The pH value were varied by adding NaOH and HCl, pH value were adjusting by using pH meter. **Table (3)** shows the samples' preparation conditions.

Table. (3):pH set solutions' preparation.

Samples	Solutions Composition	pH
S1 pH1	0.5M $NaHCO_3$+3 drops HCl	6.53
S1 pH2	0.5M $NaHCO_3$	8.06
S1 pH3	0.5M $NaHCO_3$+0.25g NaOH	9.01
S1 pH4	0.5M $NaHCO_3$+0.5g NaOH	10.43

Experiments conditions were: temp. =31 °C, soil to solution ratio=1:30, then the samples were left for 2 days.

F. Oxidative Reagent Set

Potassium Permanganate $KMnO_4$ was added to the batch solution (0.5M $NaHCO_3$) of three soil samples (0.8 g) under the conditions temperature=25 °C, soil to solution ratio=1:30) and left for 2 days of leaching, the samples and its solution were prepared as shown in **Table(4)**:

Table (4):Oxidative reagent set solution preparation.

Samples	Batch Solution
S1O1	0.5M $NaHCO_3$+0.016 g $KMnO_4$
S1O2	0.5M $NaHCO_3$+0.024 g $KMnO_4$
S1O3	0.5M $NaHCO_3$+0.032 g $KMnO_4$

It was important to study the oxidation effect on the treatment, so in the same condition but without adding $KMnO_4$, sample S1bo was prepared.

G. Soil to Solution Ratio Set

In this set five different soil to solution ratios were prepared by using (0.5M $NaHCO_3$) at 19 °C (Soil/ Solution Ratio=1:10, 1:20, 1:30, 1:40, 1:50), The samples were left for 2 days in the leaching solution.

H. Rinsing Effect

Distilled water were used to enhance the treatment process, two experiments were made; treatment with 0.5M $NaHCO_3$ followed by one rinsing (2 days left in distilled water) and by two rinsing (4 days left in distilled water), the sample before rinsing were prepared at 26 °C, with ratio 1:30, and treated with 0.5M $NaHCO_3$.

Detection of uranium in soil

A method of Solid State Nuclear Track Detectors ($SSNTD_s$) was used in the detection of depleted uranium in soil sample.

- **The Track Detectors**

Commercially available sheets of CR-39 plastic which are presently known to be the most sensitive SSNTDs and also characterized by low background were used in the present work.

These detector sheets of 250 µm thick where cut into small pieces each of (1 cm x 1 cm) area. The present sheets of CR-39 were made by Pershore modeling limited Company, U.K. The detector sheets were stored at normal laboratory conditions.

- **Experimental Procedure for uranium concentration measurement in soil**

The soil samples were prepared as previous suggestion.0.5g of soil samples were mixed with 0.1 g of methylcellulose powder ($C_6H_{10}O_5$) used as a binding material. The mixture was pressed by using a mechanical compressing device with force equal to (5 tons) in to a pellet of 1 cm diameter and 1.5mm thickness.

The pellets were covered with (CR-39) detector and placed in a plate of paraffin wax at a distance of (5 cm) from the neutron source (Am-Be), with flounce of thermal neutron (3.024 x10^9 n.cm^{-2}) and flux (5 x 1 0^3 n. $cm^{-2}.s^{-1}$), to obtain induced fission fragments from the

$$^{235}U + {}^{1}_{0}n_{thermal} \rightarrow {}^{236}U \rightarrow {}^{92}_{36}Kr + {}^{141}_{56}Ba + 3n + Q + 0.85\ Mev \quad (5)$$

After the irradiation time (7 days) , (CR-39) detectors were etched in (6.25N) NaOH solution at temperature of 60 °C for (6 h) , then the induced fission tracks density were recorded using the optical microscope.

The metal's samples were cut in small pieces and irradiated as mentioned in the above procedure for soil.

The density of fission tracks (ρ) in the samples was calculated according to the following relation [17].

Track detectors (ρ)=Average number of total pits(tracks)/Area of field view.

The uranium concentrations in soil samples were measured by comparison between track densities registed on the detectors of the sample pellet and that of the standard geological sample pellets from the relation [18,19]**(Fig.(1)):**

$$Cx(sample)/\rho x(sample)=Cs(standard)/\rho s(standard) \quad (6)$$

$$Cx=Cs.(\rho x/\rho s).$$

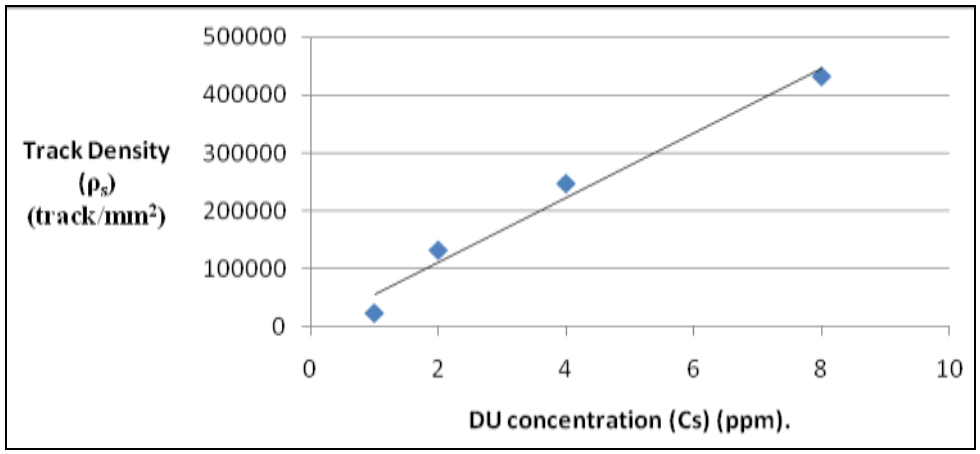


Fig.(1):Relation between standard track density and standard DU concentration[19]

Where:

Cx: uranium concentration in unknown sample (ppm).

Cs: uranium concentration in standard sample (ppm).

ρx: track density of unknown sample (track/mm^2).

ρs: track density of standard sample (track/mm^2).

RESULTS & DISCUSSIONS

Time Effect

The removal rate of uranium at various leaching time (2, 3, 4, 5 and 7days) was studied and it was found that the removal rate increased with increasing time, the concentrations of uranium after treatment were obtained between [6.903 ppm (2 days) to 4.54 ppm (7 days) with best removal efficiency equal to (63.73%) at (14 ^0C).

The concentration of DU after treatment were decreased gradually because the soil particles exposed more to solution, and that gave the chance to all the DU particles to reacts with the solution, the last two samples were approximately equal in DU removal, this was because the solution reached the saturation level and no important decrease were regarded. Relation between time and DU concentration plotted in Fig. (2).

Temperature Effect

Five soil samples with five values of temperature (20, 30, 40, 50 and 60) ^0C gave uranium concentration [from 5.857 ppm (20 ^0C) to 4.483 ppm (60 ^0C)] with best removal efficiency of DU=64.2%

It was regarded that the uranium concentrations obtained from the treatment were decreased with increasing temperature.

Temperature more than 60 ^0C was not taken, because $NaHCO_3$ will disintegrate to its ions in this temperature degree [20]. Relation between temperature and uranium concentration were plotted in Fig(3).

CARBONATE CONCENTRATION EFFECT

Five concentrations of $NaHCO_3$ (0.2, 0.3, 0.5, 0.7, and 0.8M) were examined. The resulted uranium concentration were decreased slightly with increasing sodium bicarbonate molarity [4.924 ppm at (0.2M) to 4.328 ppm at (0.8M)], the best removal efficiency of DU were equal to (65.4%).

All the previous sets and this set , the solution color after treatment was (light yellow), but for (0.8M)was different ,it had a green color, this color might due to the oxides of uranium which had different colors, and the green color referred to U_3O_8[21], the yellow color referred to UO_3 oxide. The relation between carbonate molarity and uranium concentration was plotted in fig(4).

CARBONATE/BICARBONATE RATIO EFFECT

Three samples were prepared with three different ratios of Na_2CO_3 to $NaHCO_3$, two similar results were obtained for the (S1c1/b) and (S1c/b1), the two samples had either carbonate or bicarbonate to be the dominant, and gave the same effect in treatment (about 4.6 ppm), but the sample S1c/b had equal weight of (HCO_3^{-1}) and (CO_3^{-2}), so it gave a great effect in treatment of (1.47 ppm of DU) with removal efficiency (88.25%), this difference in the treatment from S1c/b and the other two samples was due to the high concentration of (HCO_3^{-1}) and/or (CO_3^{-2}) which prevented the precipitation of uranium as hydroxide in the solution[7]. Relation between carbonate/bicarbonate and DU concentration was plotted in fig.(5). It can observe from this figure that the optimum value for Carbonate/Bicarbonate ratio equal to about 1.25 which give about 1.1ppm of Depleted Uranium. The three solutions of the samples colored with light green after treatment, and this might be resulted from the existence of (U_3O_8) in sample solution.

PH EFFECT

The effect of pH on the DU removal efficiency from soil by leaching showed in Fig. (6). (the uranium concentration versus pH values), Solutions' pH were adjusting by pH meter. The uranium removal efficiency decreased with increasing pH solution, the best treatment were in pH = 6.53(5.235 ppm of DU), with best removal efficiency of (58.2%). Where the oxidized uranium (VI) is soluble at low pH[5].

The treatment efficiency of samples set were not differed greatly, sample of (pH=9.1) had a very light green color solution after treatment, while a sample of (pH=6.53) had a yellow color.

OXIDATIVE REAGENT EFFECT

Four samples in this set were prepared; one of the samples was prepared to show the treatment efficiency without oxidation and in the same experiment conditions the other three samples were prepared (with the use of $KMnO_4$) to show the effect of oxidation on the treatment process.

The best treatment was in S1O1 (4.95 ppm) with removal efficiency of (60.4%) at (0.2gm ($KMnO_4$)/1 gm of soil) and with further increasing of $KMnO_4$ concentration the removal efficiency were decreased, this was happened because the increase of $KMnO_4$ in

the solution decrease the $KMnO_4$ capacity to solubilize the DU, and some of $KMnO_4$ were not reacted and still in the solution.

This was clearly regarded when the slight violet color appeared in the solution of S1O2, S1O3, but in S1O1 the violet color disappeared completely and the oxide gave the best treatment result in this set with the completely reaction of $KMnO_4$ in solution. Fig.(7).represented the relation between oxide and DU concentration.

SOIL /SOLUTION RATIO EFFECT

In this set five different soil/solution ratio were used. The uranium concentration in soil after treatment were widely differed from first sample and the last one (8.086ppm to 4.87ppm) and the best removal efficiency equal to (61.08%), where the concentration of DU were decreased with increasing soil/solution ratio.

This was happened because those DU particles have a great chance to react and solubilize with a high ratio of solution.

Not important notes about the solution were recorded; it had the same ordinary color of (light yellow). Relation between DU concentration and ratios were plotted in fig.(8).

RINSING EFFECT

The first sample in this set were prepared before rinsing and the two other were after rinsing, after one day and two days respectively, the best one were after two days (4.3 ppm) with removal of (65.6%).Fig(9).displayed the rinsing effect.

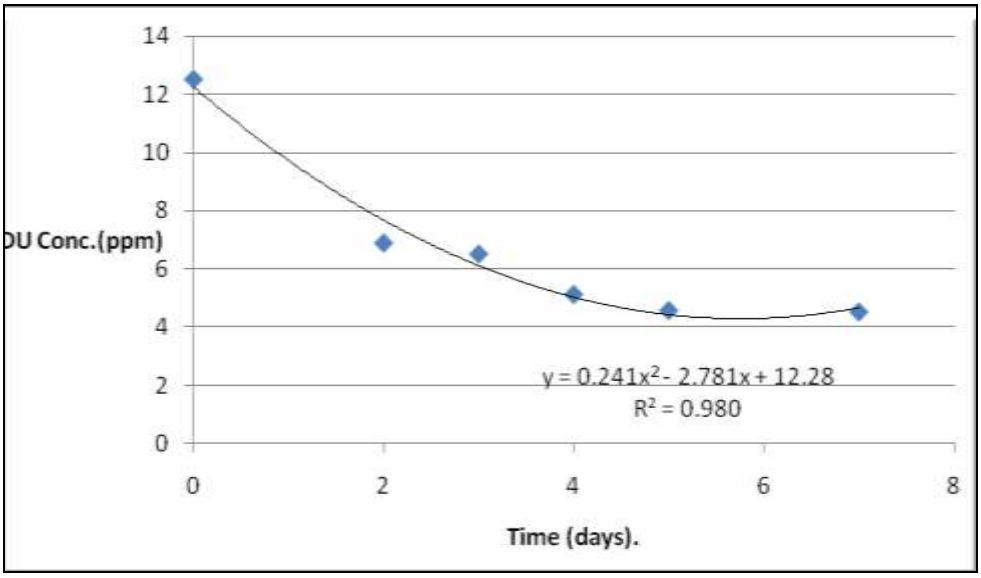


Fig.(2):Relation between DU concentration & time of leaching (Temp.=14 ^{0}C , pH=8.06 ,Solution Conc.=0.5M $NaHCO_3$,Soil:Solution=1:20).

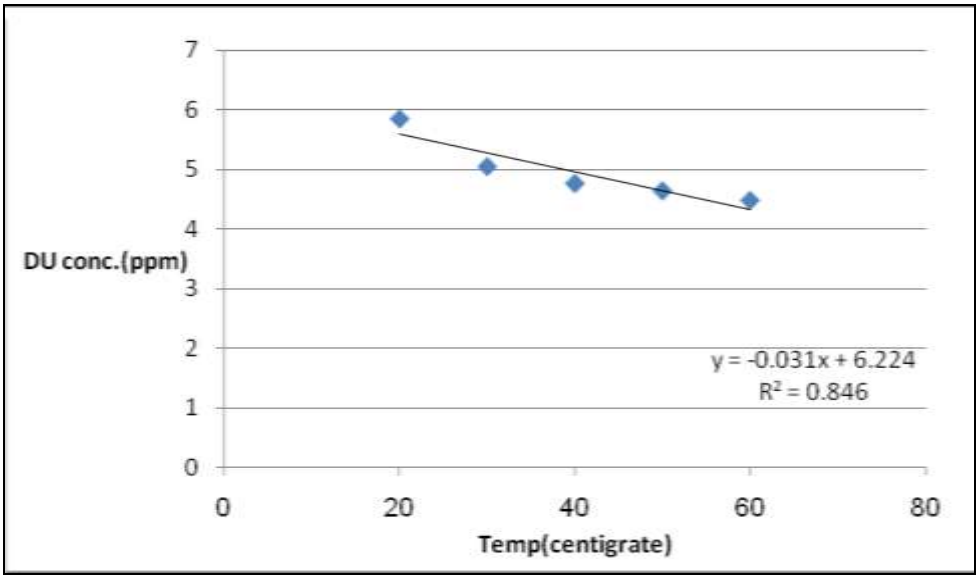


Fig.(3):Relation between DU concentration & solution's temperature. (Leaching time=3 and half hour , pH=8.06 , Solution Conc.=0.5M $NaHCO_3$,Soil:Solution=1:20).

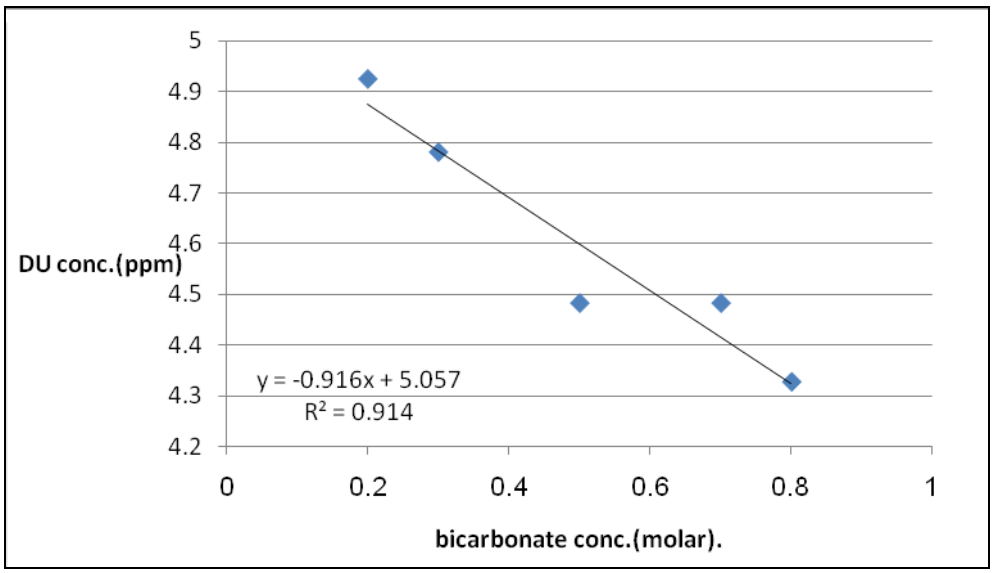


Fig.(4):Relation between DU concentration & $NaHCO_3$ concentration (Temp.=23 ^0C , pH=8.06 ,Leaching time=4 days ,Soil:Solution=1:20).

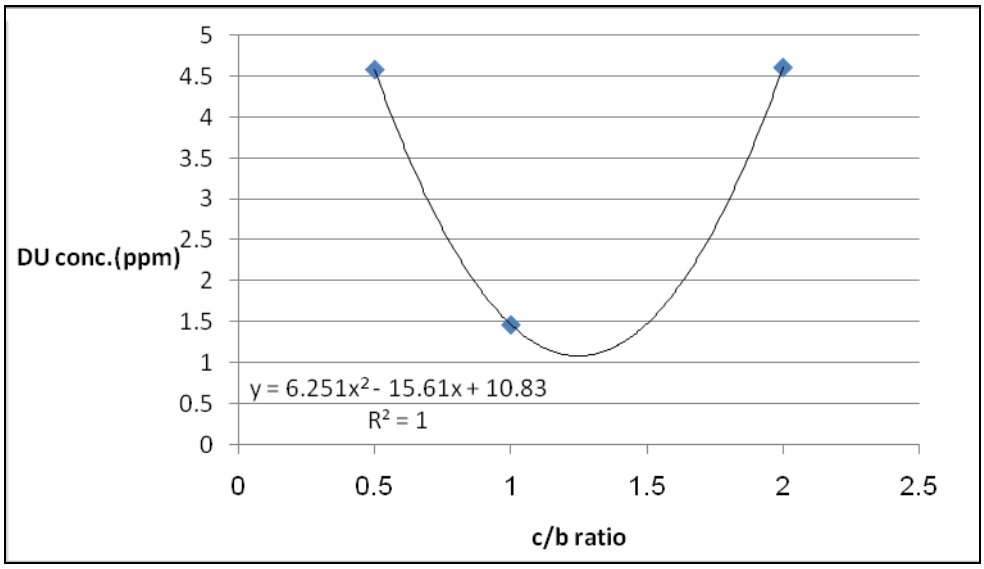


Fig.(5):Relation between DU concentration & $NaHCO_3/Na_2CO_3$ Ratio. (Temp.=26 ^0C,Leaching time=3 days , Soil: Solution=1:30).

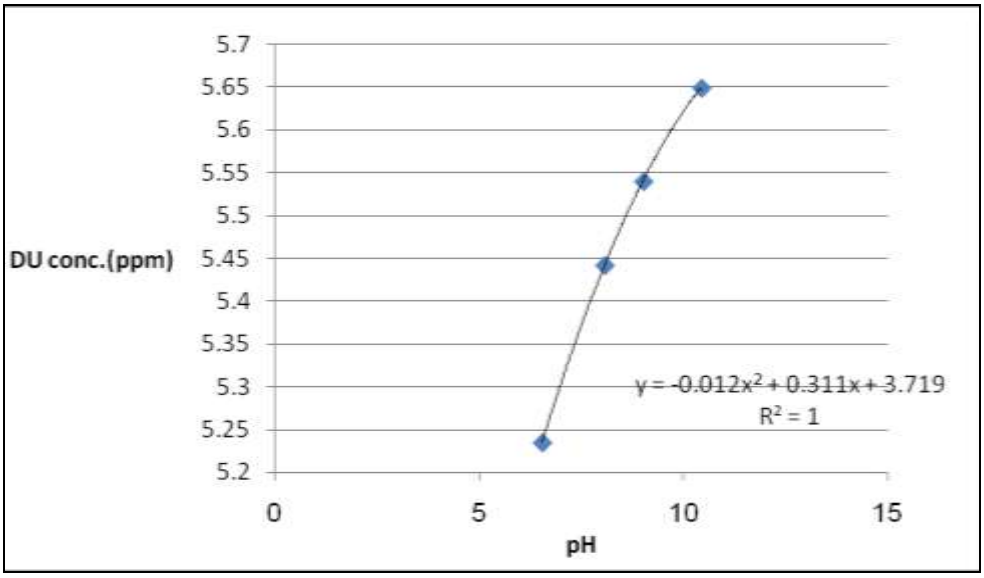


Fig.(6):Relation between DU concentration &PH of solution. (Temp.=31 ^{0}C,Leaching time=2 days , Solution Conc.=0.5M $NaHCO_3$,Soil:Solution=1:30).

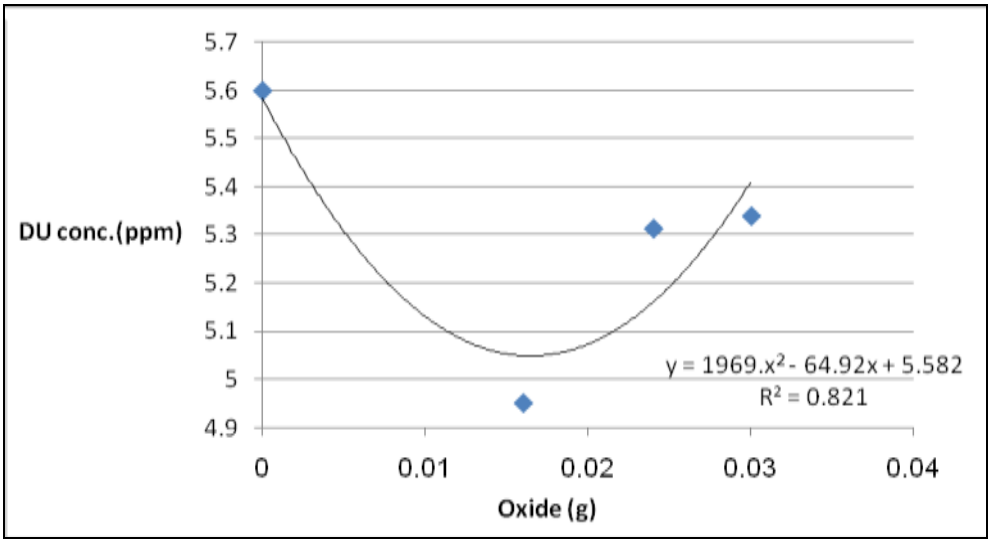


Fig.(7):Relation between DU concentration & Oxide's concentration. (Temp.=25 ^{0}C,Leaching time=2 days , solution conc.=0.5M $NaHCO_3$,soil:solution=1:30).

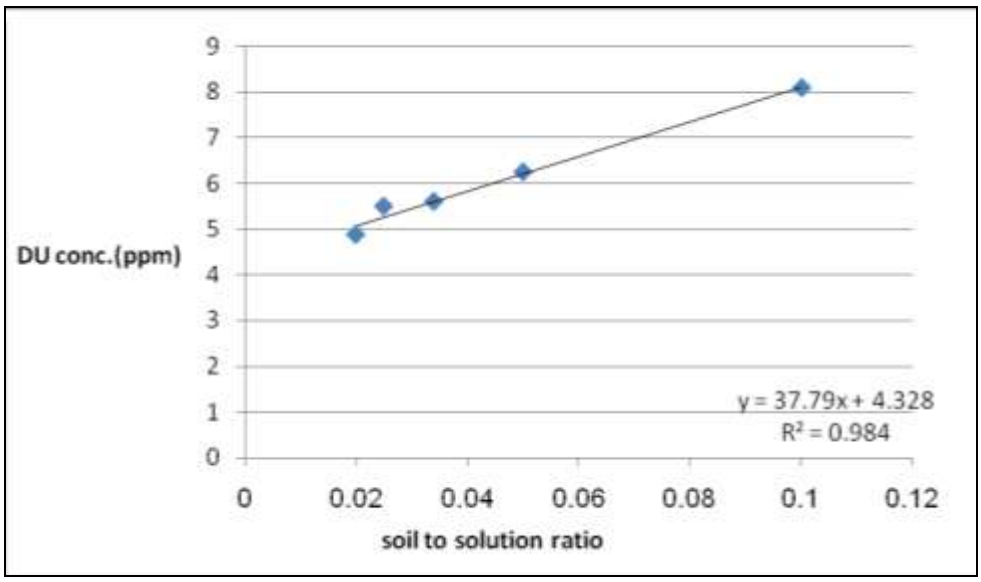


Fig.(8):Relation between DU concentration & soil to solution ratio.

(Temp.=19 ^{0}C,Leaching time=2 days , Solution Conc.=0.5M $NaHCO_3$).

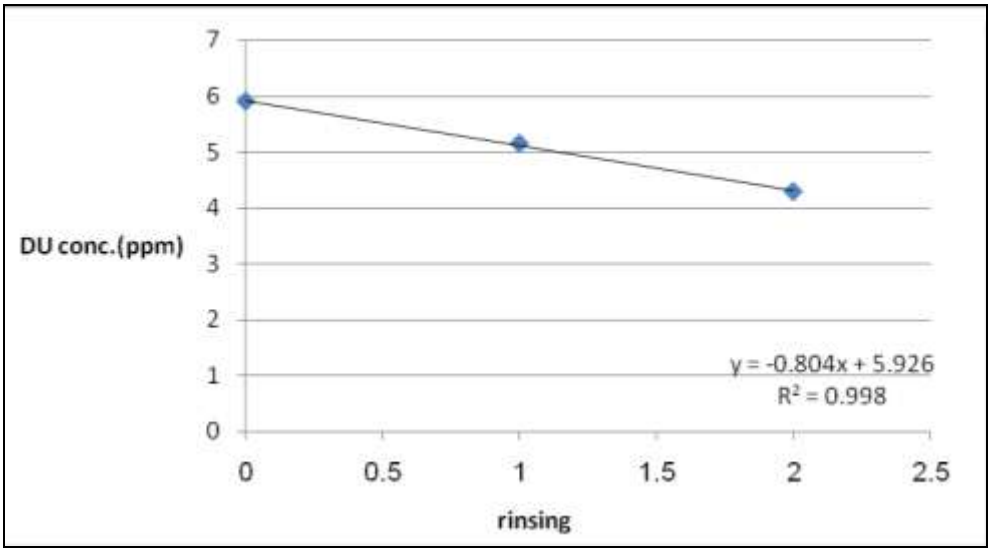


Fig.(9):Relation between DU concentration & Rinsing after treatment.

(Temp.=26 ^{0}C,Leaching time with 0.5M $NaHCO_3$ =2 days, Soil:Solution=1:30).

Conclusions

a. with increasing leaching time the DU concentration in soil after treatment were decreased. Removal efficiency were ranged between (44.8-63.7)%
b. With increasing solution temperature the DU concentration after treatment in soil were decreased. Removal efficiency were ranged between (53.2-64.2)%
c. With increasing Bicarbonate concentration, the DU concentration in soil after treatment were decreased. Removal efficiency were ranged between (60.6-65.4)%
d. Carbonate/bicarbonate ratio: the best treatment was at ratio=1, when the weight of carbonate were equal to bicarbonate. Removal efficiency were ranged between (63.1-88.25)%.that was good result to decrease the DU concentration lower than the allowed level of 40 Bq/Kg or about 1.6 ppm.
e. With increasing pH the DU concentration in soil were also increased. Removal efficiency were ranged between (60.8-54.8)%
f. The influence of oxidizing reagent were decreased as the increasing of its weight added to solution, the best treatment were at oxidizing ratio to soil(0.02g KMnO4/1g of soil). Removal efficiency were increased from (2.3 to 5.2)% from before oxidation.
g. With increasing the soil/solution ratio the DU concentration in soil were decreased. Removal efficiency were ranged between (35.4 - 61.08)%
h. Taking two conditions of rinsing with distilled water giving good effect after treatment with bicarbonate leaching. Removal efficiency increased from (6 to12.8) % from before rinsing.

References

- United Nation Environmental Program, UNEP (2003); "***Depleted Uranium in Bosnia and Herzegovina***" Post Conflict Environmental Assessment.
- Chester W. Francis, Michael E.Timpson, Suk Y.Lee, Mark P.Elles and Jim H. Wilson , ,(1997);"***The use of carbonate lixiviants to remove uranium from uranium-contaminated soils."*** Environmental sciences division, Oak Ridge National Laboratory
- E.H.P. CORDFUNKE, ***(***1969);"***The Chemistry of Uranium, including its applications in nuclear technology"*.**
- M.C.DUFF and C. AMRHEIN. (1996); "***Uranium(VI) Adsorption on Goethite and Soil in Carbonate Solutions,*** "Soil Sci. Soc. Amer. J, 60,1393-1400.

- C. W. Francis ,J. R. Brainard, D. J. Chaiko, Gretchen Matthern, and D. A. York (1994);"***Uranium Removal From Soils: An Overview From The Uranium in Soils Integrated Demonstration Program.***
- D.L CLARK, D.E. HOBART, and M.P. NEU.(1995);"***Actinide Carbonate Complexes and Their Importance in Actinide Environmental Chemistry***,"Chem. Rev. 95,25-48
- C.F.V. Mason.; W.R.J.R. Turney; B.M.Thomson; N.LU; P.A.Longmire and C.J.Chisholm-Brause,(1997);"***Carbonate leaching of uranium from contaminated soil***".
- Martine C. Duff, Caroline F. V. Mason, J. A. Musgrave, (1997);"***Comparison of Acid and Base Leach For The Removal of Uranium From Contaminated Media***"
- C. W. Francis, A. J. Mattus, M. P. Elless, and M. E. Timpson,(1993); " ***Carbonate-and Citrate-Based Selective Leaching of Uranium-Contaminated Soils***" Part 1,In"Removal of Uranium from Uranium-Contaminated Soils, Phase I: Bench-Scale Testing."ORNL-6762. Oak Ridge National Laboratory
- A. J. Mattus, C. W. Francis, L. L. Farr, M. P. Elless and S. Y. Lee, (1993);" ***Selective Leaching of Uranium From Uranium-Contaminated Soils***".
- M. E. Timpson, M. P. Elless, andC. W. Francis,(1994);" ***Influence of Attrition Scrubbing, Ultrasonic Treatment, and Oxidant Additions on Uranium Removal from Contaminated Soils***".
- M. P. Elless, S. Y. Lee and M. E. Timpson, (1993); "***Physicochemical and Mineralogical Characterization of Uranium-Contaminated Soils From The Fernald Integrated Demonstration Site***".
- Leung, K.C.; Lau, S.Y.; Poon, C.B (1990); "***Gamma Radiation Dose From Radionuclides in Hong Kong Soil***"; J.Environ Radioactivity;11; 279-290; Elsevier Science Publishers Ltd
-] Lin, Y.; Lin, P.; Chen, C.; Haung, C. (June,1987);"***Measurement of Terrestrial & Radiation in Taiwan, Republic of China***" Health Physics; Vol.52; No.6; 805-811; Health Physics Society.
- Mollah, A.S.; Rahman, M.M.; Husain,S.R. (June,1986). "***Distribution of γ-emitting Radionuclides in Soils at the Atomic Energy Research Establishment, Saver, Bangladesh***" Health Physics; Vol.50; No.6; 835-838;Health Physics Society.
- Kataoka, T.; Ikebe, Y.; Minato, S.; Ishida, K.; (June,1982) "***Detailed Evaluation of Natural Gamma-Radiation Field due to Uranium(U-238) Series***"; Journal of Nuclear Science and Technology; Vol.19; No.6; 482-490
-] Amald, O., Custball, N.H and Nielsen, G.A. (1989), "***Cs-137in Montarq Soils***", Health Physics, 57 No.6, 955-958.
- Fleischer, R.L. , Price, P.B. and Walker, R.M. ,(1975)." ***Nuclear Tracks in Solids: Principles and Applications***", Pergammon press, Berkeley
-] Bansal, V.,Azam, A. and Prasad, R. (1989)"***Health Physics***", Vol. 27, PP 985.
- http://www.mineralszone.com
- http://www.webelements.com

Sets	Sample name	Description.
Time set	S1a	Sample 1 at time=2 days
	S1b	Sample 1 at time=3 days
	S1c	Sample 1 at time=4 days
	S1d	Sample 1 at time=5 days
	S1e	Sample 1 at time=7 days
Temperature set	S1t1	Sample 1 at temp.=20 C^o
	S1t2	Sample 1at temp.=30 C^o
	S1t3	Sample 1at temp.=40 C^o
	S1t4	Sample 1at temp.=50 C^o
	S1t5	Sample 1at temp.=60 C^o
$NaHCO_3$ concentration set	S1k1	Sample 1 at M=0.2
	S1k2	Sample 1 at M=0.3
	S1k3	Sample 1 at M=0.5
	S1k4	Sample 1 at M=0.7
	S1k5	Sample 1 at M=0.8
Carbonate/Bicarbonate Ratio set	S1c1/b	Sample 1at $Na2CO_3>NaHCO_3$
	S1c/b1	Sample 1at $Na2CO_3<NaHCO_3$
	S1c/b	Sample 1at $Na2CO_3=NaHCO_3$
pH set	S1pH1	Sample 1at pH=6.53
	S1pH2	Sample 1at pH=9.01
	S1pH3	Sample 1at pH=8.06
	S1pH4	Sample 1at pH=10.43
Oxidative reagent set	S1bo	Sample 1 before oxidation
	S1o1	Sample 1 at KMnO4=0.02/1g soil
	S1o2	Sample 1 at KMnO4=0.03/1g soil
	S1o3	Sample 1 at KMnO4=0.04/1g soil
Soil/Solution ratio	S1r1	Sample 1at soil/solution ratio=1:10
	S1r2	Sample 1at soil/solution ratio=1:20
	S1r3	Sample 1at soil/solution ratio=1:30
	S1r4	Sample 1at soil/solution ratio=1:40
	S1r5	Sample 1at soil/solution ratio=1:50
Rinsing Effect	S1br	Sample 1before rinsing
	S1rr1	Sample 1 after one rinsing
	S1rr2	Sample 1 after two rinsing
Other Solution	S1ca	Sample 1 with citric acid
	S1sa	Sample 1 with sulfuric acid

	S1ci/bi	Sample 1with citric acid then $NaHCO_3$

NUMERICAL INVESTIGATION OF NATURAL CONVECTION IN A VERTICAL ANNULUS ENCLOSURE

Ass. Prof. Manal H. AL-Hafidh
University of Baghdad
Mechanical Eng. Dept

Safa Bontok Raheem
University of Baghdad
Mechanical Eng. Dept

ABSTRACT

A numerical technique is developed to predict both the transient and steady axisymmetric two-dimensional natural convection heat transfer for water as the working fluid in a vertical annulus enclosure of a fixed radius ratio (2) aspect ratio (1) and Rayleigh number ranging within ($10^3 \leq Ra_d \leq 10^6$) for a fixed Prandtl number (Pr=7). Finite difference analogs of the Navier – Stokes and thermal energy equations are solved in the stream function – vorticity frame work. The results obtained are presented graphically in the form of streamline, vorticity and isotherm contour plots. A correlation has been set up to give the average Nusselt number variation with Ra_d and for which the results are found to be in good agreement with previously published experimental data.

الخلاصة

يتضمن البحث دراسة رقمية باستخدام طريقة الفروق المحددة لدراسة انتقال الطاقة الحرارية بالحمل الحر لماء في حيّز حلقي عمودي ثنائي البعد لأسطوانتين متمركزتين مغلقتي النهايات ولنسبة أقطار ثابتة(٢) لنسبة باعية [As =ι/ $(r_o - r_i)$] تساوي (١) وخواص جريان ($10^3 \leq Ra_d \leq 10^6$) و(Pr=7). تم حل معادلات الاستمرارية والزخم والطاقة الانتقالية بعد تحليلها إلى معادلات لا بعدية ومن ثَمّ إلى دالّة الانسياب الدوّاميّة. النتائج تم توضيحها برسم منحنيات مغلقة للانسياب ومنحنيات درجات الحرارة الثابتة، والنتائج أعطت تمثيلاً لمعدل رقم نسلت ($\tilde{N}u_i$) ضد Ra_d. المقارنة مع البحوث الأخرى أعطت نتائج دقيقة كفاية لإمكانية حساب معامل انتقال الطاقة الحرارية لحالة الجريان المذكورة آنفاً وللماء كوسط لانتقال الحرارة .

KEY WORDS: Natural Convection, Concentric Vertical Annulus, Laminar Flow, Numerical Solution

INTRODUCTION

The annulus represents a common geometry employed in a variety of heat transfer systems ranging from simple heat exchangers to the most complicated nuclear reactors. In spite of the importance of convection heat transfer in vertical annular enclosures in many practical applications, very few basic studies have so far been conducted for this system. Many finite difference solutions of free convection problems of long horizontal, rectangular and cylindrical enclosures subject to lateral heating are found in the works of **(Boyd, 1983), (Charrier-Mojtabi and Mojtabi et-al, 1979), (Date, 1986), (Kuhen & Golstein, 1976) and (Akbar et-al, 1985).** Numerical studies of free convection in vertical annulus enclosures are found in the works of **(Schwab & De Witt, 1970), (Kubair & Simha, 1982) (Keyhani et-al, 1983) and (Prasad & Kulacki, 1985)**. Many investigators interested in the study of eccentric annular enclosures **(Shue et-al, 2001)** or elliptic cylinders **(Lee and Lee, 1981).** Most of the studies available are for high aspect ratio.

The present paper considers the axisymmetric flow regime in a vertical annulus enclosure **Fig. (1)** whose surface temperatures are kept isothermal and with Boussinesq approximation being made to the governing equations.

MATHEMATICAL FORMULATION

The problem considered are both the transient and steady state, two-dimensional axisymmetric laminar convection of a Boussinesq fluid initially at rest and a uniform temperature is assumed in the mathematical model for the system shown in **Fig. (1).** The vertical cylindrical surfaces are considered to be perfect conductors of heat with the inner wall temperature (Ti) greater than that of the outer wall (To). Both top and bottom surfaces of the enclosure are considered to be perfect insulators with a rigid and motionless bounding surfaces. The dimensionless equations of continuity, momentum and energy are **(Schwab & De Witt, 1970)**:

$$({}^{1}/_{R})[R\,(\partial U/\partial R)] + (\partial V/\partial Z) = 0 \qquad (1)$$

$$(\partial U/\partial T_M)+U(\partial U/\partial R)+V(\partial U/\partial Z) = -(\partial P/\partial R)+Pr[\partial/\partial R\{({}^{1}/_{R})(\partial RU/\partial R)\}+(\partial^2 V/\partial Z^2)] \qquad (2)$$

$$(\partial V/\partial T_M)+V(\partial V/\partial R)+(\partial V/\partial Z) = -(\partial P/\partial Z)+Pr[({}^{1}/_{R})(\partial/\partial R\{R(\partial V/\partial R)\}+(\partial^2 V/\partial Z^2)] \qquad (3)$$

$$(\partial \Theta/\partial T_M)+U\,(\partial \Theta/\partial R)+V(\partial \Theta/\partial Z) = ({}^{1}/_{R})[(\partial/\partial R)\{R(\partial \Theta/\partial R)\}]+(\partial^2 \Theta/\partial Z^2) \qquad (4)$$

The vorticity-stream function formulation is applied in order to avoid direct determination of the lateral variations of pressure. The conservation equations then become:

Vorticity:

$$(\partial \omega/\partial T_M) + (\partial U\omega/\partial R) + (\partial V\omega/\partial Z) = Ra_d\, Pr\,(\partial \Theta/\partial R) + Pr\,[(\partial/\partial R)\ \{({}^{1}/_{R})\ (\partial R\omega/\partial R)\} \qquad (5)$$

Stream function:

$$\nabla^2\psi = -\omega = (^1/_R)\,[(\partial^2\psi/\partial R^2) - (^1/_R)\,(\partial\psi/\partial R) + (\partial^2\psi/\partial Z^2)] \quad (6)$$

Energy:

$$(^1/_R)(\partial RU\Theta/\partial R) + (\partial V\Theta/\partial Z) = (^1/_R)\,[(\partial/\partial R)\{R(\partial\Theta/\partial R)\}] + (\partial^2\Theta/\partial Z^2) \quad (7)$$

The initial and boundary conditions are:

For $T_M = 0$, $\omega = \psi = U = V = 0$ (no slip condition)

, $\Theta_i = 1$, $\Theta_o = 1$ (constant wall temperature)

For $T_M > 0$, $\psi = \partial\psi/\partial n = 0$, $\partial\Theta/\partial Z = 0$

$$\omega_{\text{Vertical wall}} = -(^1/_R)\,(\partial^2\psi/Z^2) \quad , \quad \omega_{\text{horizontal wall}} = -(^1/_Z)(\partial^2\psi/R^2)$$

NUMERICAL SOLUTION OF THE GOVERNING EQUATIONS

An analytic solution cannot be found for a set of equations because of the complexity of those characterizing the cavity problem, and thus they must be integrated numerically by finite difference techniques. In this study, the governing equations which are expressions of conservation of mass, momentum and thermal energy are non-dimensionalized. A two–dimensional explicit finite difference technique was used to solve the transient behavior of the fluid and the heat transfer until the steady state is reached by marching out in time step ΔT_M. A forward difference technique may used to convert the convection terms in the energy and vorticity transport equations to an algebraic terms and in concern of diffusion terms, the central difference technique may used. To get numerical stability, the forward difference technique is applied when the mean value of the velocity is positive and the backward difference technique is applied when the mean velocity value is negative, so the method choice depends on the flow direction. To solve the stream function equation the Gauss Siedel method used with the successive over relaxation method to make earlier convergence. A numerical stability occurred when the time step ΔT_M is greater than $(\Delta T_M)_{max.}$, which depends on Ra_d and grid spacing. In this study the suitable ΔT_M was found to be ($\Delta T_M \leq 0.0001$) for fixed Ra_d.
The main steps of the solution procedure can be listed as follow:

1. The boundary and initial conditions are specified for the dependent variables.
2. The discretized temperature equation is solved to obtain the updated temperature field.
3. Vorticity at all interior grid points was similarly advanced from using the updated temperature field.
4. Stream function at all interior grid points was updated with the updated vorticity field.
5. Local Nusselt number computed and the method of successive over relaxation was employed to obtain the new stream function at each grid point. Convergence was assumed whenever any two successive iterations was less than 10^{-4}.

6. Total heat transfer calculated depending on local Nusselt number.
7. Steps (1-6) were repeated until the magnitude of the error ratio (ε) did not change by more than 10^{-4}.

RESULTS AND DISCUSSION

The difference equations of this study were solved on a digital computer using FORTRAN 95 program and the Tec Plot program used to plot the isotherm and streamline contours. The results are expressed in the transient region until reaching steady state where $\tilde{N}u_i$ will be constant with time.

Transient Results

Grid Spacing

The time steps used depend on Ra_d as shown in **Table (1)**, and the grid spacing which was found to give an adequate representation of the results for $Ra_d=10^3$ and 10^4 was (1/10) and for $Ra_d=10^5$ and 10^6 the grid spacing were (1/20) and (1/30) respectively.

Table (2) explain the best values for nodes number (N) in radial direction and (M) in vertical direction for ($10^3 \leq Ra_d \leq 10^6$).

Flow and Isotherm Patterns

Figs. (2-13) show typical results obtained for vertical annuli with Rayleigh number ($10^3 \leq Ra_d \leq 10^6$), for different time steps. The fluid near the outer (colder) cylinder is heavier and is moving downward while the relatively lighter fluid near the inner (hotter) cylinder is moving upward and with the increase of Ra_d a slightly vertical displaced occurred towards the top of the annulus so the convection regime appears clearly as shown in **Figs. (2, 3, 4 and 5).**

The temperature gradient across the cavity is horizontal with conduction profiles being vertical lines traversing the entire length of the cavity as indicated in **Fig.(6)** and then start to develop and indicate more inclination in its lines while the slope of lines near (cold wall) is small because the velocity is decreased **Fig.(7, 8 and 9)**. When Ra_d approaches 10^3, large temperature gradients grow near the vertical wall giving rise to the formation of thermal boundary layers and fluid velocities sufficient to form hydrodynamic boundary layer. A unicellular flow pattern is generated in the enclosure. It is noticed that the maximum heat transfer, indicated by closely spaced isotherms, is located at the top of the cavity for the outer cylinder and at the bottom for the inner one.

As a consequence of the symmetry and the continuity, the resulting fluid motion inside the half cavity consists of one rotating vortex as shown in **Fig. (10)** when the time increasing a multi small vorticity will be appeared

In **Figs. (11, 12 and 13)** it is seen that a small secondary vorticity appears in the bottom and become smaller with time increasing where a start of forming of a secondary vorticity near the outer cylinder will happened which will be grow with the increasing of fluid velocity.

Temperature Distribution and Local Nusselt Number

Fig. (14) shows the temperature variation versus Ra_d for a gap width (z = 0.5). In the thermal boundary layer adjacent to the outer cylinder, it is seen that the temperature gradient increases considerably and the trend is opposite near the inner cylinder. This is probably due to higher heat transfer by angular convection flow rather than radial one.

A correlation of ($\tilde{N}u_i$) with respect of Ra_d was set up which is given as:

$$\tilde{N}u_i = 0.4134 * Ra_d^{0.287} \quad (8)$$

Steady State Results

Heat transfer rate can be calculated at a steady state when the average Nusselt number ($\tilde{N}u_i$), reach a constant value with time as shown in **Fig. (15).**

Steady state streamlines and the isotherms for Ra_d ranging within (10^3- 10^6) are explained in **Fig. (16 a & b) and Fig.(17 a & b)** respectively and the vorticity lines are explained in **Fig.(18 a& b)** which shows the multi-cellular flow regime except at ($Ra_d > 10^5$), where single vorticity appears.

A reasonable agreement between the present work results and the previous researches are shown in **Fig. (19).**

Table (1) Time Steps used with different Ra_d

ΔTM	Ra_d
0.0001	10^3
0.0001	10^4
0.00005	10^5
0.00001	10^6

Table (2) Best Values for the Selected Node Number for Study Cases

Ra_d	N	M
10^3- 10^4	11	11
10^5	21	21
10^6	31	31

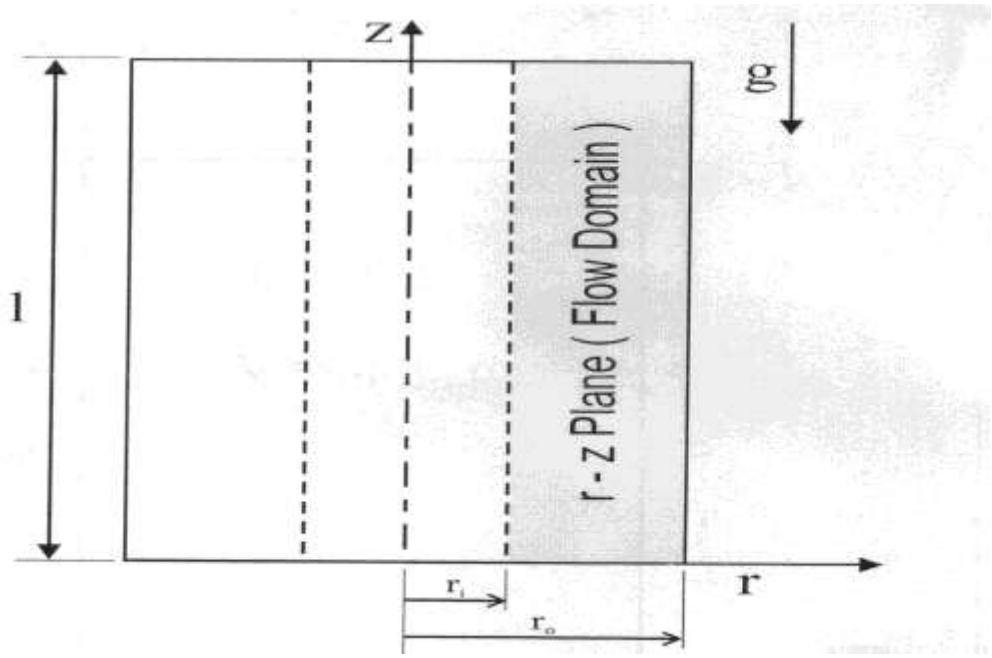


Fig. (1) Enclosure Geometry And Coordinate System

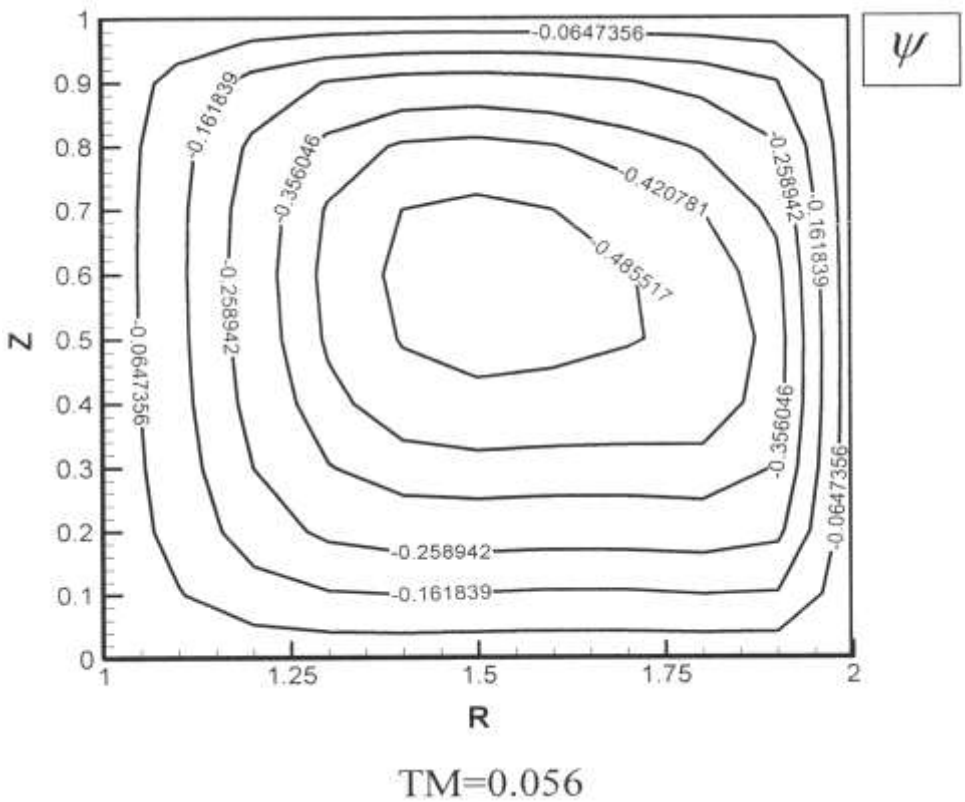
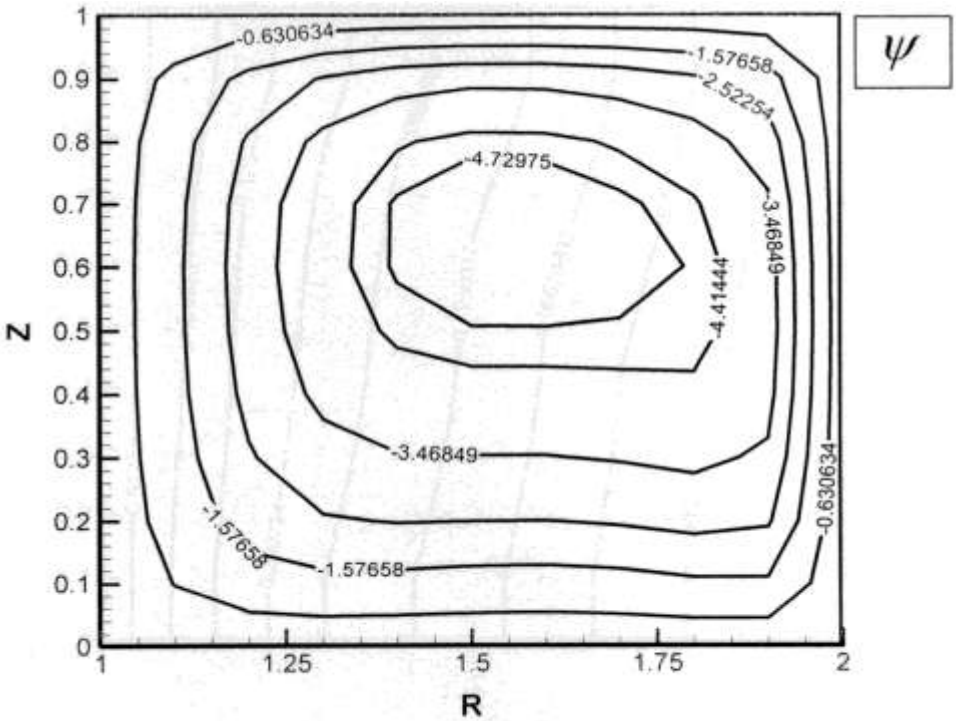
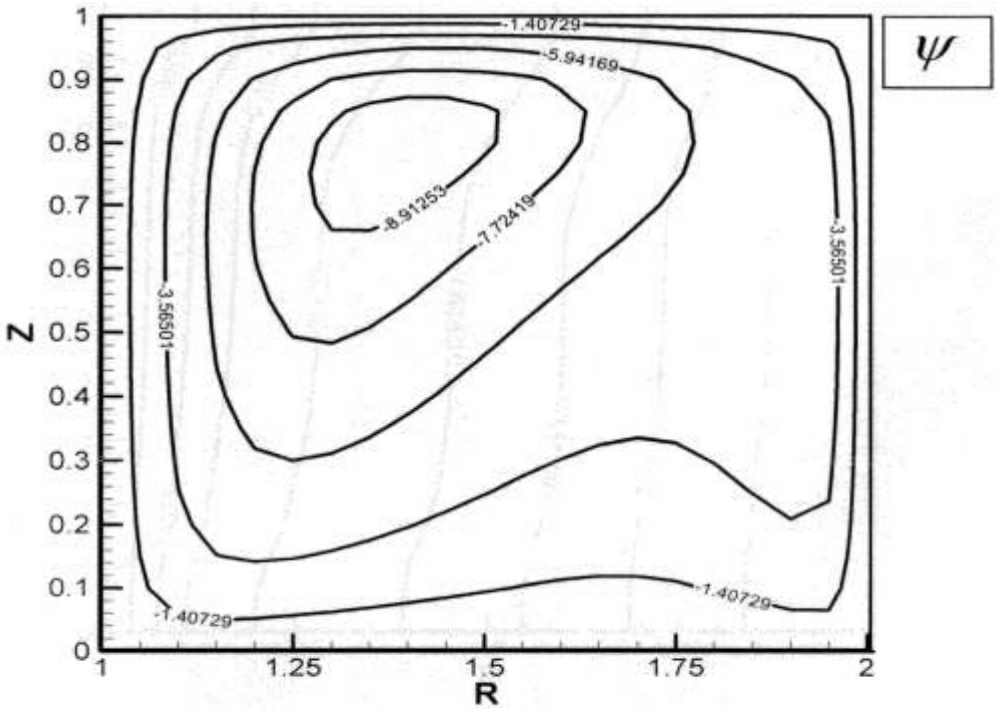


Fig. (2) Transient Streamlines for $Ra_d = 10^3$



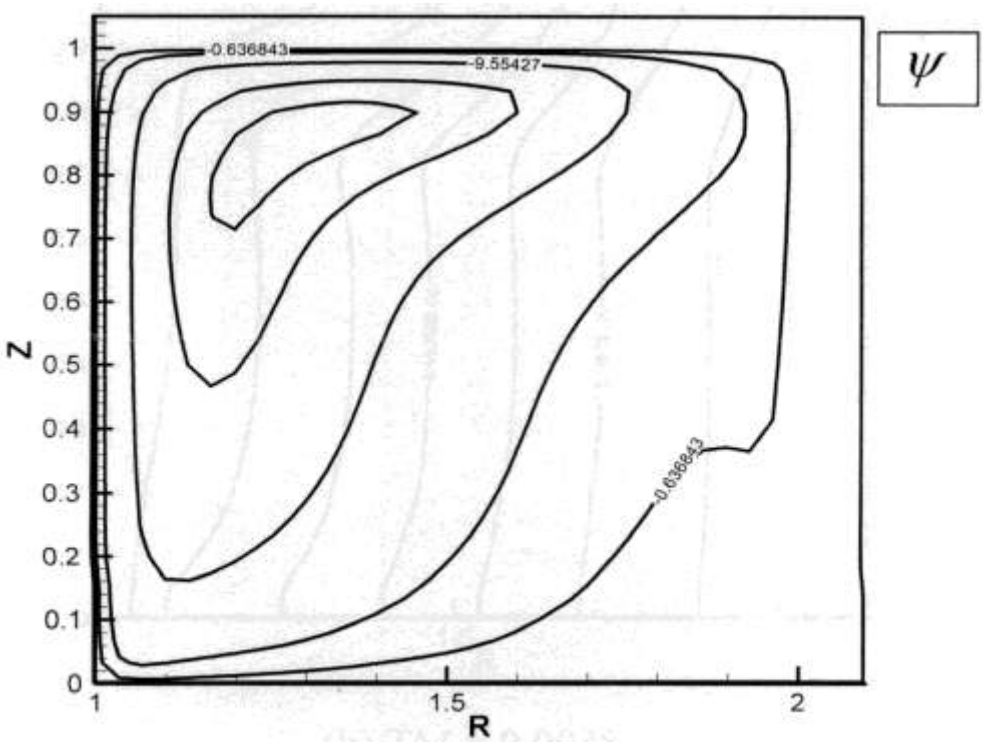
TM = 0.048

Fig. (3) Transient Streamlines for $Ra_d = 10^4$



TM = 0.024

Fig. (4) Transient Streamlines for $Ra_d = 10^5$



TM = 0.0094

Fig. (5) Transient Streamlines for $Ra_d = 10^6$

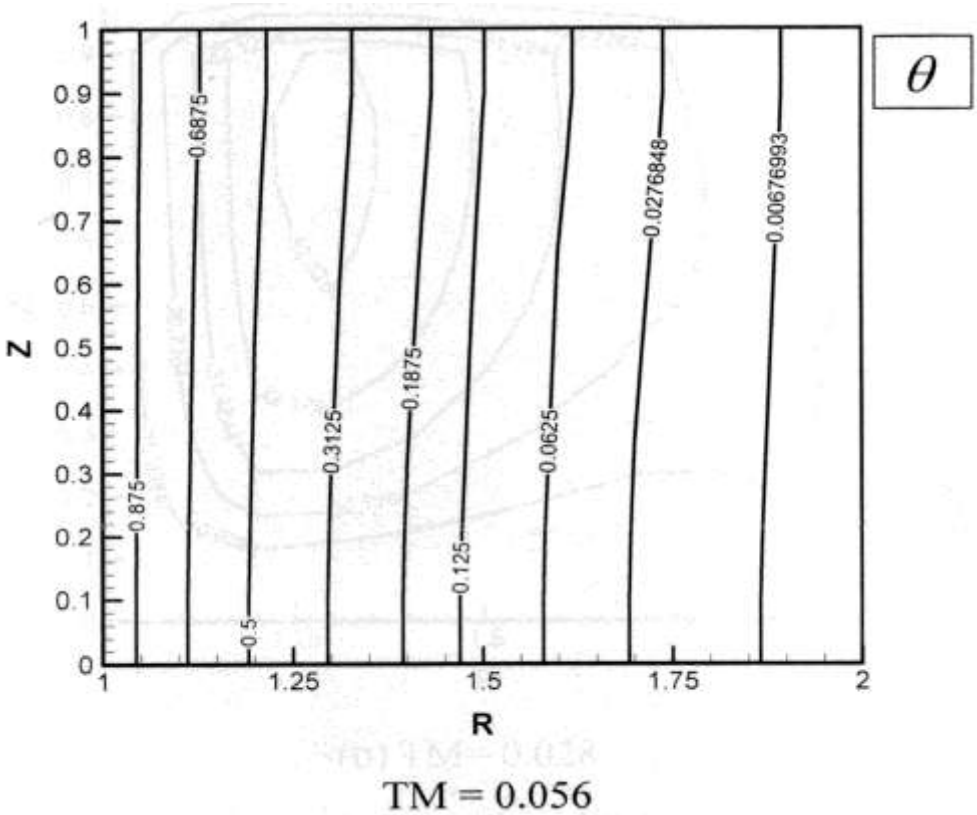


Fig. (6) Transient Isotherms for $Ra_d = 10^3$

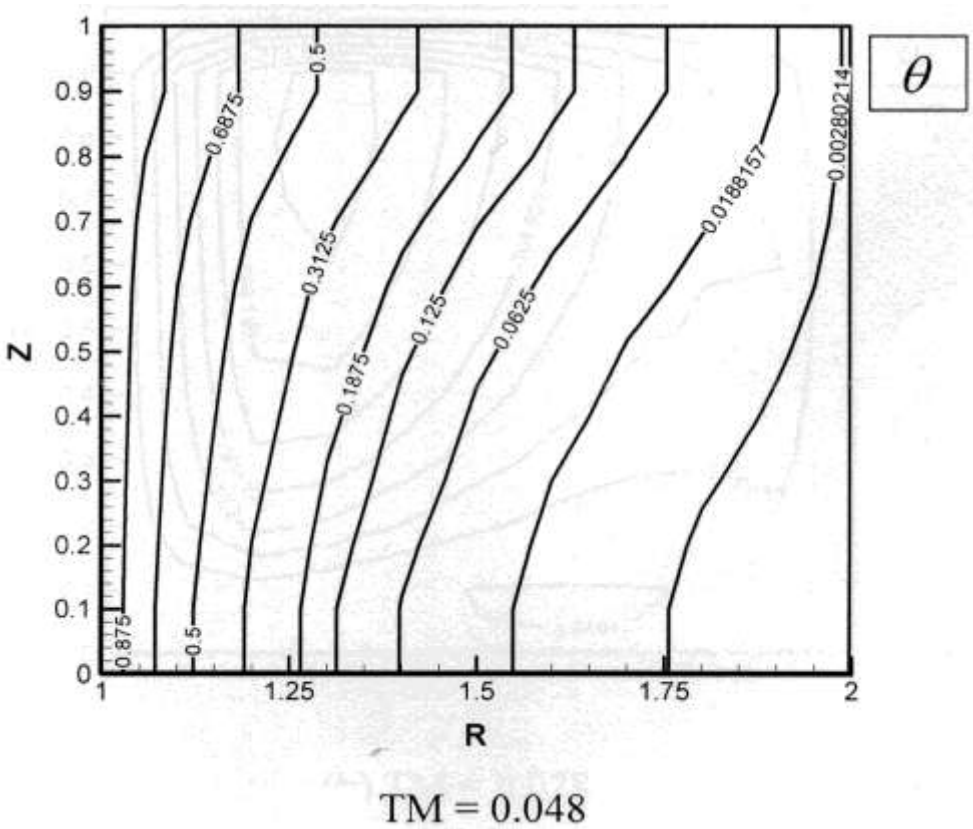


Fig. (7) Transient Isotherms for $Ra_d = 10^4$

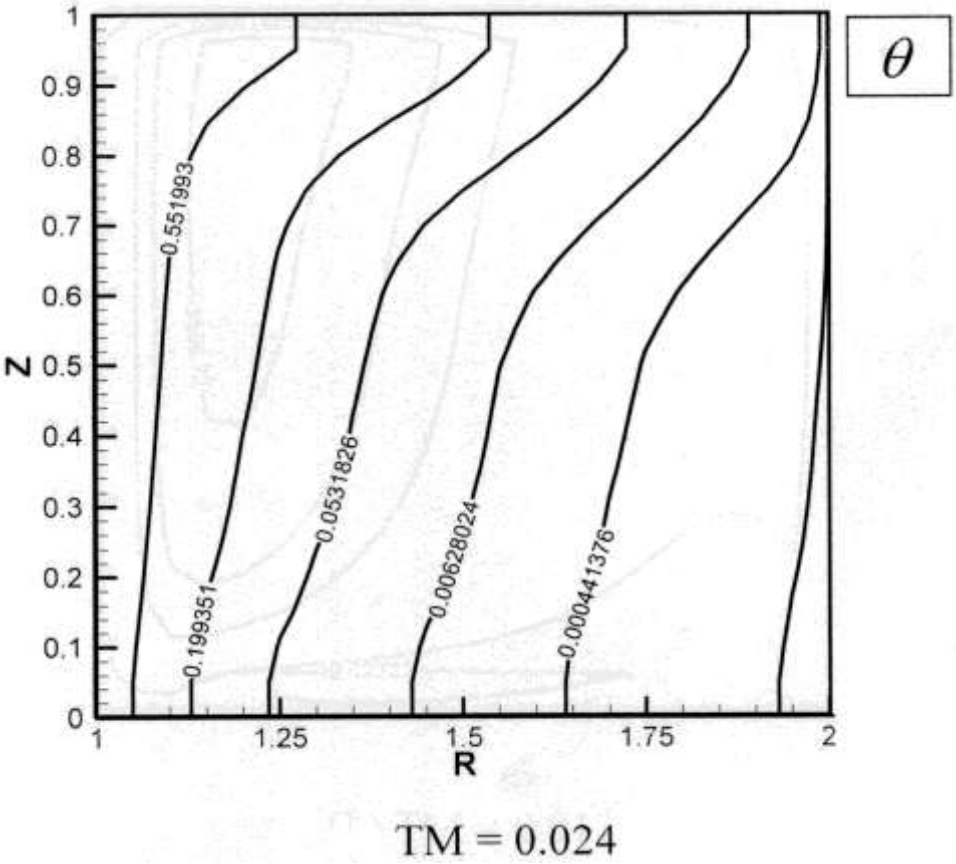


Fig. (8) Transient Isotherms for $Ra_d = 10^5$

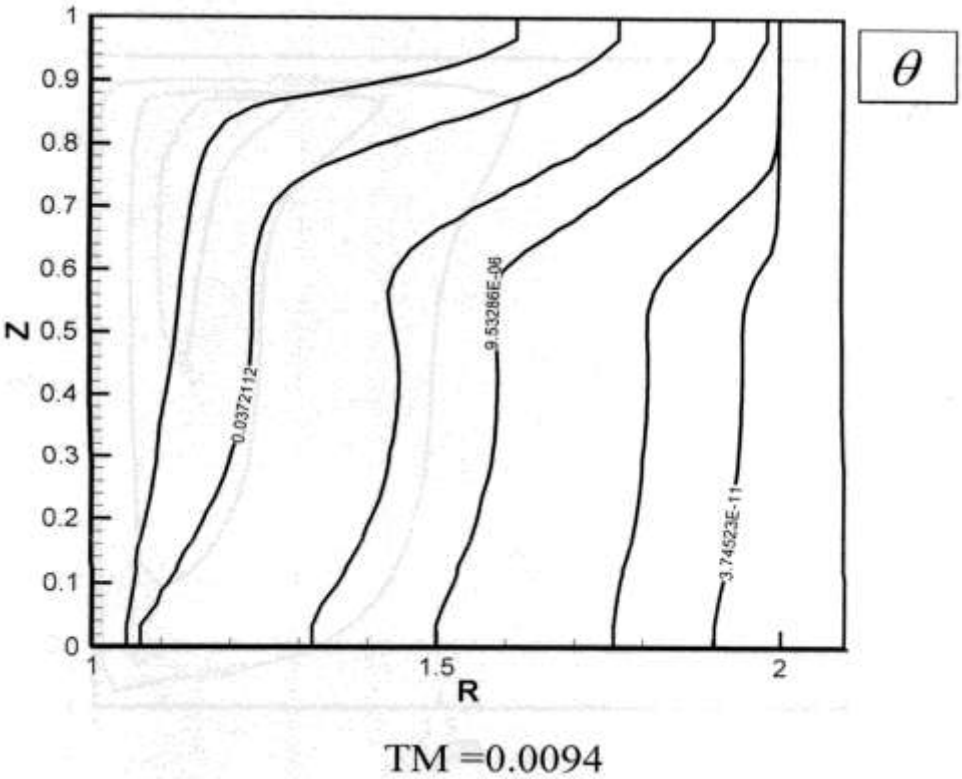


Fig. (9) Transient Isotherms for $Ra_d = 10^6$

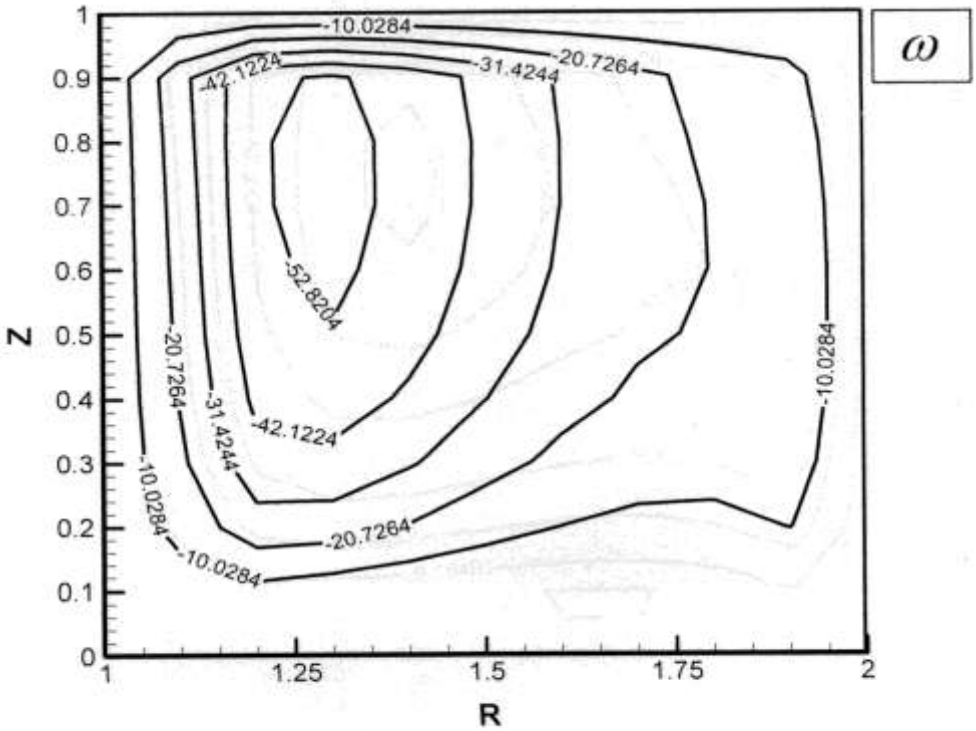


Fig. (10) Transient Vorticity for $Ra_d = 10^3$

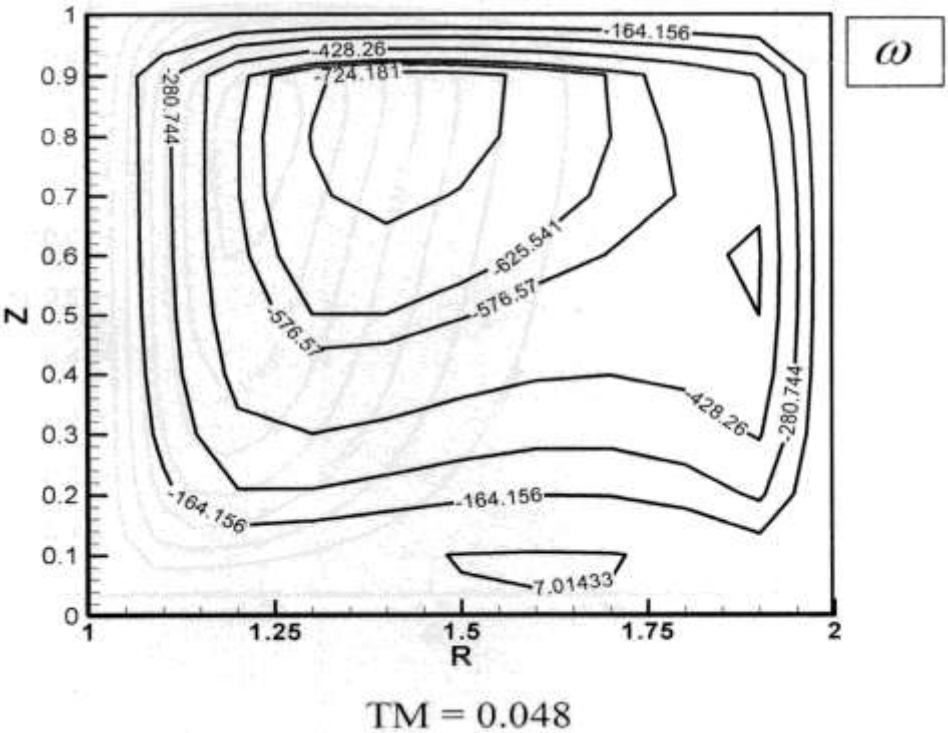


Fig. (11) Transient Vorticity for $Ra_d = 10^4$

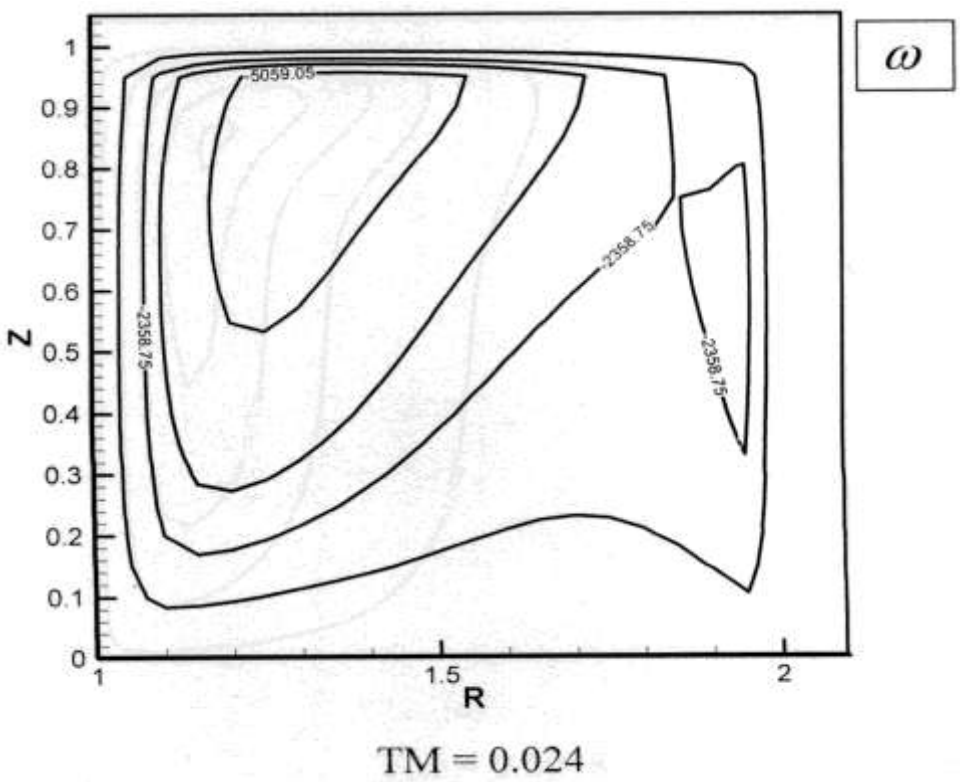


Fig. (12) Transient Vorticity for $Ra_d = 10^5$

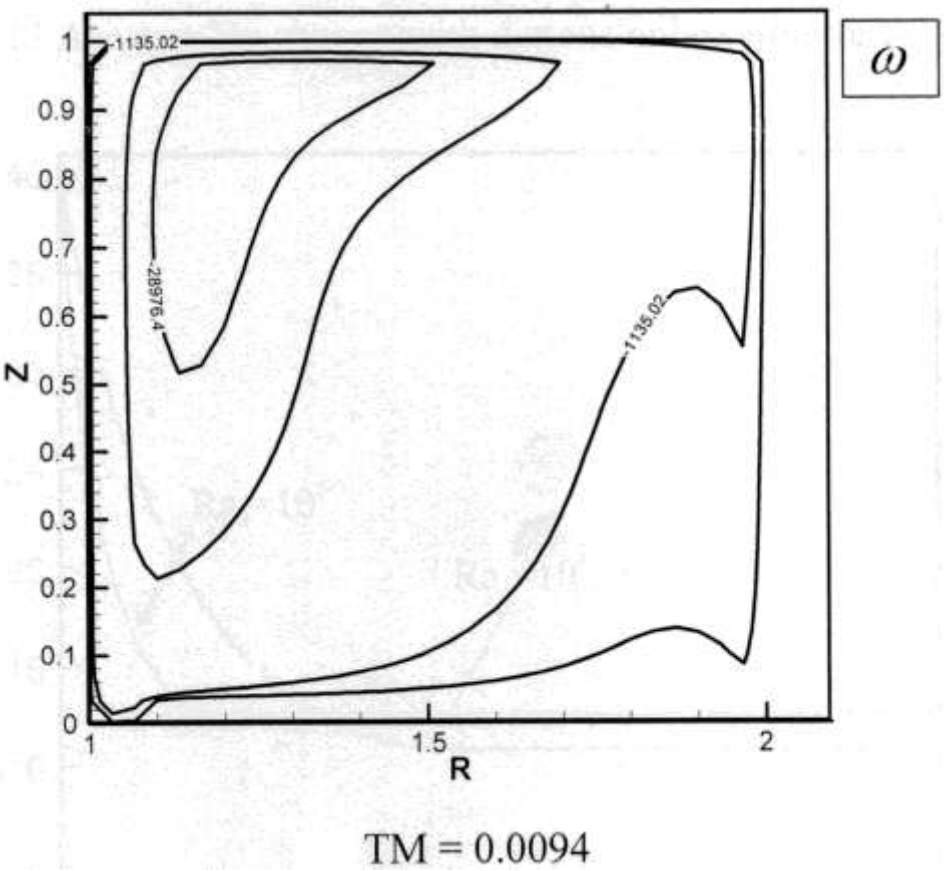


Fig. (13) Transient Vorticity for $Ra_d = 10^6$

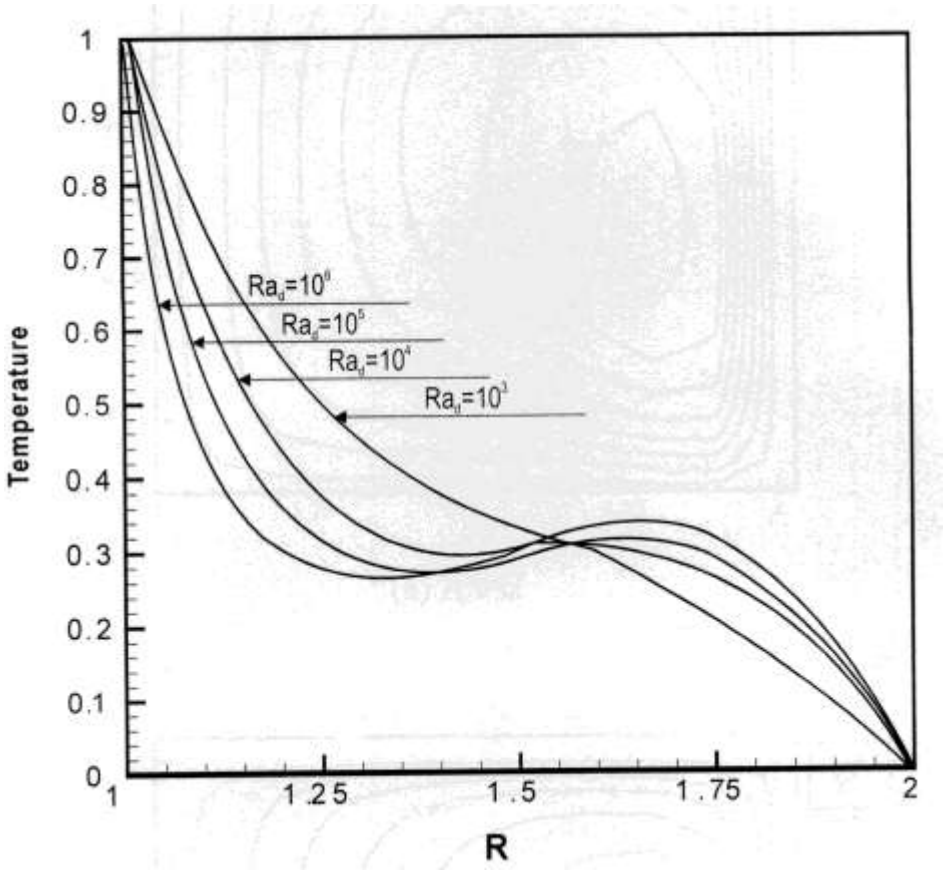


Fig. (14) Temperature Distribution with Gap Width at Z = 0.5

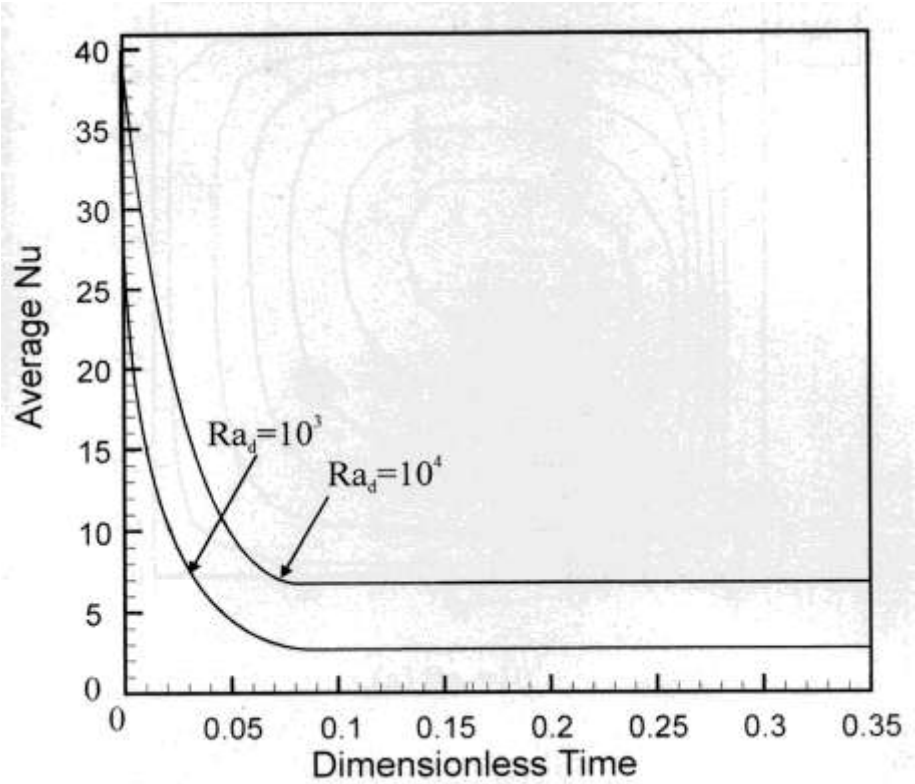


Fig. (15) Average Nu_i Change with Dimensionless Time for $Ra_d = 10^3$ & 10^4

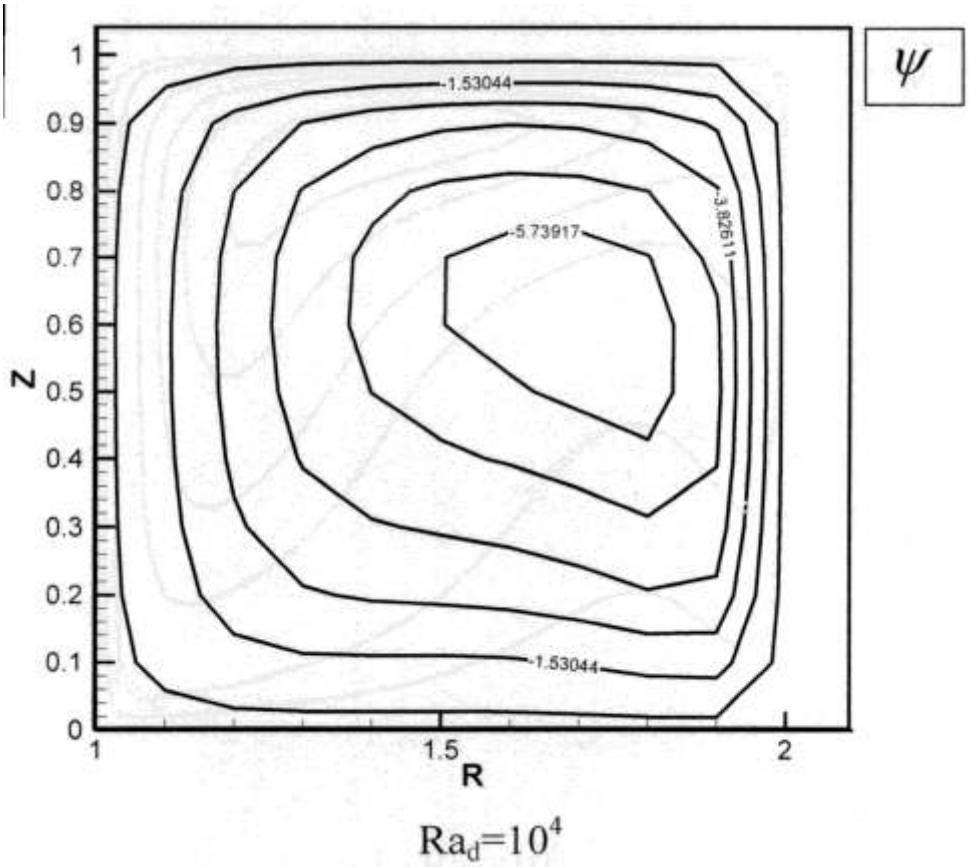


Fig. (16a) Steady State Streamlines for $Ra_d = 10^4$

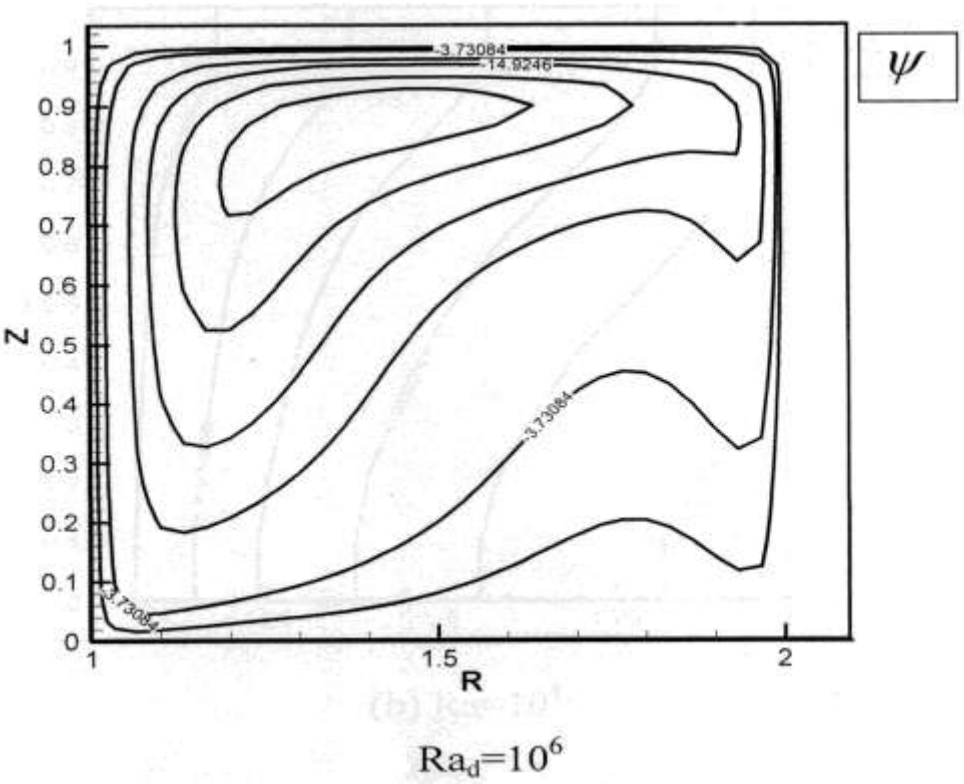


Fig. (16b) Steady State Streamlines for $Ra_d = 10^6$

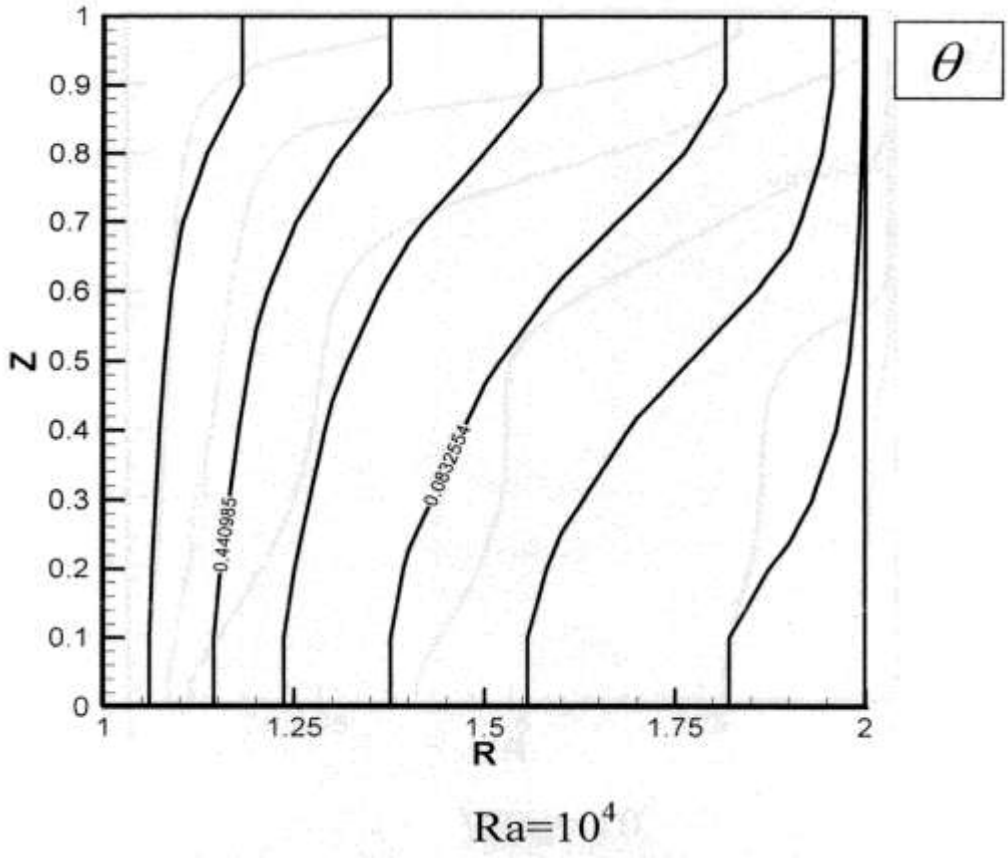
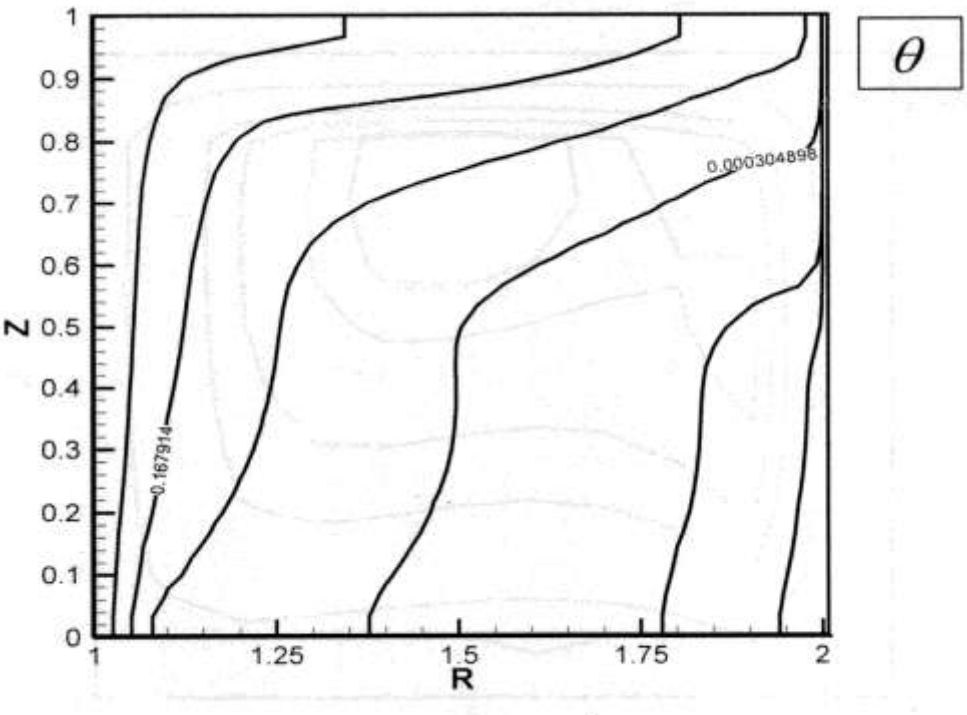
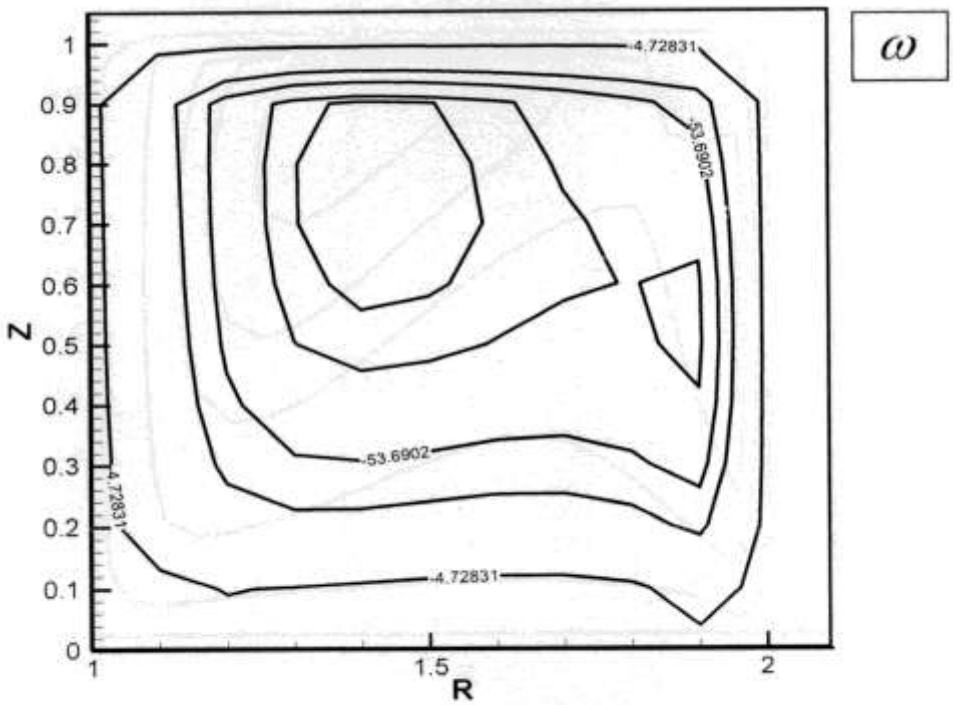


Fig. (17a) Steady State Isotherms for $Ra_d = 10^4$



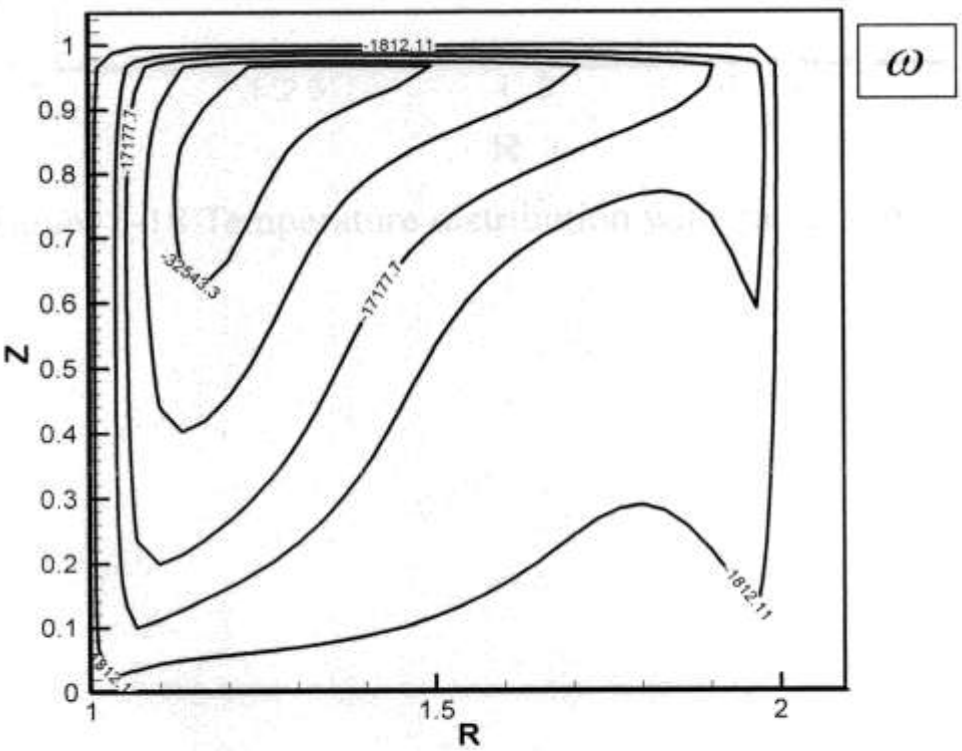
$Ra_d = 10^6$

Fig. (17b) Steady State Isotherms for $Ra_d = 10^6$



$Ra_d = 10^3$

Fig. (18a) Steady State Vorticity Lines for $Ra_d = 10^3$



$Ra_d=10^6$

Fig. (18b) Steady State Vorticity Lines for $Ra_d = 10^6$

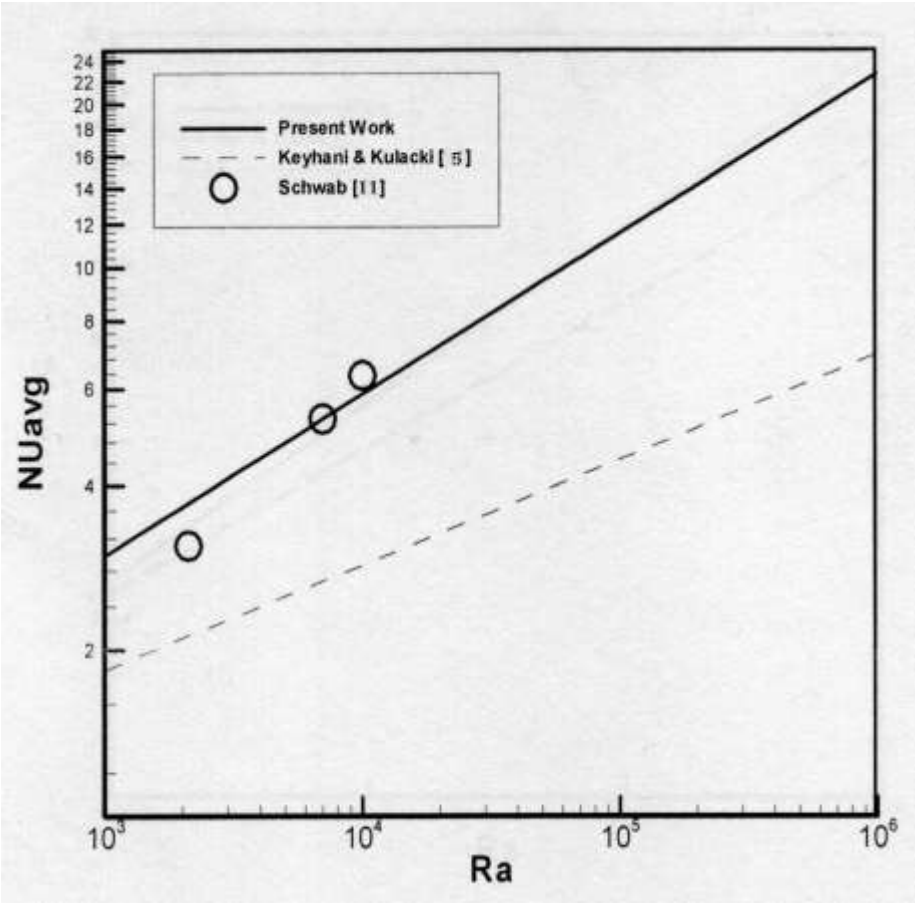


Fig. (19) Comparison between Present Work and Keyhani [5] and Schwab [11]

CONCLUSIONS

The natural convection of a mass of water contained between two concentric cylinders has been investigated numerically.

The results obtained in the present study may be summarized as follows:

1. The basic operation for heat transfer is conduction for ($Ra_d \leq 10^3$) and convection for ($Ra_d > 10^3$).
2. For the unsteady state the streamlines shows a single cell form except at ($Ra_d = 10^5$) where it shows a bicellular form.
3. The time of reaching steady state decreasing with increase of Ra_d as shown in **Fig. (15)** because the increasing of Ra_d cause an increase of liquid acceleration according to buoyancy force increasing.
4. **Fig. (17)** shows that the heat transfer operation can be divided into three regions, the first is that of heat transfer by conduction which extended to ($Ra_d=10^3$) where ($\tilde{N}u_i = 3$) and it is considered to be the start of convection region. The second region is transition region until plume region is appeared and it's ranged is ($10^3 \leq Ra_d \leq 10^4$). Third region is the plume region where ($Ra_d \geq 10^5$).

REFERENCES

Akbar Hessami, M. A., Pollard, A., Rowe, R. D., and Ruth, D. W., (1985), "A study of free convection heat transfer in horizontal annulus with a large radii ratio", ASME J. of Heat Transfer, Vol. 107, PP. 603-609.

Boyd, R., (1983), "A unified theory for correlating steady laminar natural convection heat transfer data for horizontal annuli", ASME J. of heat transfer, Vol. 105, PP. 1545-1548.

Charrier-Mojtabi, M. C., Mojtabi, A. and Caltaglrone, J. P., (1979), "Numerical solution of flow due to natural convection in horizontal cylindrical annulus", ASME J. of heat transfer, Vol. 101, PP. 171-173.

Date, A. W., (1986), "Numerical prediction of natural convection heat transfer in horizontal annulus", Int. J. of Heat and Mass Transfer, Vol. 29, PP. 1457-1464.

Keyhani M., Kulacki F. A. and Christensen R. N., (1983), "Free convection in vertical annulus with constant heat flux on inner wall", ASME J. of Heat Transfer, Vol. 105, PP. 454-459.

Kubair, V. B., and Simha, C. R. V., (1982), "Free convection heat transfer to mercury in vertical annulus", Int. J. of Heat and Mass Transfer, Vol. 25, No. 3, PP. 399-402.

Kuhen, T. H., and Golstein, R.J. (1976), "An experimental and theoretical study of natural convection in the annulus between horizontal concentric cylinders", J. of Fluid Mechanics, Vol. 74, Part 4, PP. 695-699.

Kuhen, T. H., and Golstein, R.J. (1980), "A parametric study of Prandtl number and diameter ratio effects on natural convection heat transfer in horizontal cylindrical annuli", Trans. of ASME, Vol. 102, PP. 768-770.

Lee, J. H., and Lee, T. S., (1981), "Natural convection in annuli between horizontal confocal elliptic cylinders", Int. J. of Heat and Mass Transfer, Vol. 24, No. 10, PP. 1734-1742.

Prasad, V. and Kulacki, F. A., (1985), "Free convection heat transfer liquid filled vertical annulus", ASME J. of Heat Transfer, Vol. 107, PP. 596-602.

Schwab, T. H., and De Witt, K. J., (1970), "Numerical investigation of free convection between two vertical coaxial cylinders", AICHE J., Vol. 16, PP. 1005-1010.

Shue, C., Xue, H. and Zhu, Y. D., (2001), "Numerical study of natural convection in an eccentric annulus between a square outer cylinder and circular inner cylinder using DQ method", Int. J. of Heat and Mass Transfer, Vol. 44, PP. 3321-3333.

NOMENCLATURE

Symbols	Description	Units
As	Dimensionless aspect ratio $As = \iota / (r_o - r_i)$	
g	Acceleration of gravity	m^2/sec
L	Cylinder length	m
M	Nodes number in r-direction	
N	Nodes number in z-direction	
$\tilde{N}u_i$	Average Nusselt number $\tilde{N}u_i = h (r_o - r_i)/k$	
Pr	Prandtl number $Pr = \nu/\alpha$	
R	Dimensionless radial direction $R = r/(r_o - r_i)$	
Ra_d	Rayleigh number $Ra_d = Pr (g\beta(T_i - T_o) (r_o - r_i)^3/\nu^2)$	
r_i, r_o	Inner and outer radius	m
T_M	Dimensionless time $T_M = \alpha t / (r_o - r_i)^2$	
t	Time	sec
T_f	Film temperature $T_f = (T_i + T_o) /2$	K
U	Dimensionless radial velocity $U = u (r_o - r_i) /\alpha$	
u	Radial velocity	m/sec
V	Dimensionless vertical velocity $V = v (r_o - r_i) /\alpha$	
v	Vertical velocity	m/sec
Z	Dimensionless vertical direction $Z = z / (r_o - r_i)$	
z	Gap width $z = (r_o - r_i)$	m

GREEK SYMBOLS

Symbols	Description	Units
ω	Dimensionless vorticity	
ψ	Dimensionless stream function	
ε	Error ratio	
Θ	Dimensionless temperature $\Theta = (T - T_o) / (T_i - T_o)$	
α	Thermal diffusivity	m^2/sec
β	Volume coefficient of expansion $\beta = 1/T_f$	1/K
ν	Kinematics' viscosity	m^2/sec

DESIGN AND IMPLEMENTING A HEAT PIPE EXPPERIMANTAL SYSTEM FOR RESIDENTIAL HEATING

Zeina A. Al-Saadi Dr. Mikdam M. saleh

University of Baghdad-College of Eng.
Department of Nuclear Engineering

ABSTRACT:

The design and construction of an instrumented heat pipe H.P for domestic heating that uses water as a working fluid was undertaken to investigate experimentally the performance of the H.P under various operating conditions of: Power levels, water inventories and angle of inclinations. A theoretical model to predict the temperature of the condenser surface (the temperature at which heat is rejected) and system pressure at steady state conditions was developed and used to compare these parameters with the experimental findings. The model utilizes the total heat supplied to the evaporator to predict system pressure and condenser temperature. The theoretical model is suitable for vertical H.P (i.e. θ =0) and its predictions of condenser surface temperature is within ±16% and of system pressure is within ±21%. An acceptable H.P design may have a condenser heat flux of 1.16 k W/m^2 with a corresponding system pressure of 1500 kPa.

الخلاصة

تم تصميم و بناء منظومة أنبوب حراري للتدفئة المنزلية التي يستخدم الماء كمائع تشغيلي. أخذنا بنظر الاعتبار عند البحث التجريبي لأداء الأنبوب الحراري التغيرات في العوامل التشغيلية التالية:مختلف مستويات القدرة الداخلة إلى المسخن ومستويات مختلفة للمائع داخل المنظومة(Inventory) إضافة إلى زوايا ميل مختلفة (Inclination Angle). النموذج النظري الذي تم تطويره يتمكن من حساب المعدل لدرجة الحرارة السطحية للمكثف (t_c) و ضغط المنظومة(P)في حالة الاستقرار .وتمت مقارنة النتائج العملية مع الحسابات النظرية ووجد أن النموذج النظري يستخدم الحرارة الكلية المسلطة على المسخن لتوقع ضغط المنظومة ودرجة حرارة المكثف ويعتبر هذا النموذج مناسب للأنبوب الحراري الوضع العمودي (θ= ◦0) والنموذج الرياضي يعطي قيم

بحدود (±15%) لدرجة الحرارة السطحية للمكثف و(±21%) لضغط المنظومة. نستنتج أن التصميم العملي المقبول لمنظومة التدفئة المنزلية يجب أن يكون ضمن حدود كثافة فيض حراري(1.16 kw/m^2) خلال ضغط منظومة (1500 kpa) وزاوية ميل تتراوح بين(0−30 °).

KEY WORDS:
Thermosyphon Design, Natural Convection, Heat Pipe

INTRODUCTION

A heat pipe (H.P) is a heat transfer device, which has high effective thermal conductivity. Heat pipes are evacuated vessels, typically circular in cross section, which are back-filled with a small quantity of working fluid. They are totally passive and are used to transfer heat from a heat source to a heat sink with minimal temperature gradients **(Qian and Chen 1999)**. Some applications of heat pipes are in cooling electronic devices, heat recovery from exhaust ventilation, application in cooling a computer processor **(Shunji and Suzuki 2000)**, Geological applications **(Torrance 1979)** and **(Chi 1976)** , Electronics cooling **(Marquet. and Solecki 1989)**, nuclear thermonic space power concept using rod control and heat pipes **(Dunn and . Reay 1976).** Most of the published research dealt with the theory and performance of wick type heat pipe, since most of its applications are in zero gravity environment (cooling of re-entry vehicles). To the authors knowledge there has been no experimental or theoretical work done on heat pipes specifically intended for residential heating in the open literature.

The wickless H.P to be considered in this work is like a thermosyphon. The working fluid of choice was water because of its availability and low cost and also because one can use cast iron pipe for the H.P.This choice of material allows high-pressure operation of the H.P (up to 40 bars). Of course we can use copper for construction materials because of its compatibility, corrosion wise with water.

The concept of using a H.P for domestic heating stem from the idea of providing very clean heating without resorting to more expensive electrical heating. And also avoid using any moving parts. The condenser must be cooled by natural convection driven by density differences only. Fins are to be used to dissipate the rejected heat from the condenser in order to lower the thermal resistance across the condenser.

There is no application specifically tailored for residential heating using a heat pipe in the open literature. What prompt this work is the convenience of the H.P in providing large heat flux at small temperature gradient coupled with the attractive feature of using conventional and available fuel like gasoline to heat residential homes.

The present work will attempt to address the following goals:

1. Design and implement an experimental H.P rig with a finned condenser cooled by natural convection. To study its performance under a variety of operating conditions of input power to the evaporator, water inventory and angle of inclination.
2. Develop a theoretical model to predict the operation characteristics of the H.P: mainly system pressure and condenser temperature, from a single parameter, namely the total heat supplied to the evaporator.
3. Compare the theoretical predictions with experimental results.

EXPERIMENTAL APPARATUS

The rig used for this purpose was designed and built here and is shown schematically in **Fig.A** consists of a cast iron pipe of length (2.5m) which is divided into an evaporator, condenser and adiabatic zone. Carbon steel was chosen because of its compatibilty corrosion with water. Detail description of these components along with its associated instrumentation is given below.

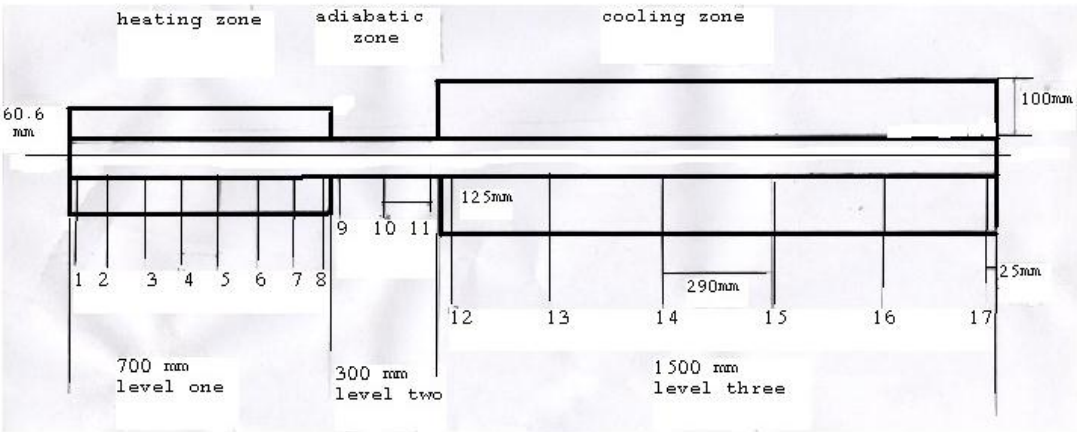


Fig.A: Experimental Rig of the H.P (Showing Thermocouple Locations)

The Evaporator:

The evaporator has a length of (0.7m) and is heated electrically by eight co-axial type heaters of (750 watt) each. The heaters were fixed to the outside surface of the evaporator at (45) degree intervals. The length of each is equal to the length of the evaporator. In order to reduce heat losses to the outside and to insure uniform heat flux along the evaporator, a heat shield in the form of a cast iron pipe of length (0.7m) and thickness (0.003m), was installed concentrically around the evaporator pipe. The heat shield is then wrapped with (0.03m) of glass wool insulator to further reduce the heat losses. In order to measure the average temperature of the evaporator, nine thermocouples were imbedded on the outside surface of the evaporator in (0.002m) deep groves. These groves were filled with lead-zinc alloy to insure good thermal contact. Three thermocouples were attached to the outside surface of the insulator to get the average temperature of that surface. All thermocouples were of (0.2mm)-asbestos sheathed Alumel-Chromel (type K).

The Condenser:

The condenser has a length of (1.5m) and was cooled by longitudinal fins that reject heat with the environment through natural convection. The fins were welded to the out side surface of the condenser. The length of each is equal to the length of the condenser with width (0.1m) and thickness (0.002m). In order to measure the average temperature of the condenser; six thermocouples were imbedded on the outside surface of the condenser. **Table.1** shows the dimensions of the fin.

The Adiabatic Zone:

The segment of the rig between the finned condenser and the evaporator is normally referred to as the adiabatic zone. The temperature profile is monitored by three thermocouples type K. **Table.2** shows the heat pipe specifications.

Table .1 fins Specifications

1	length	L=1.5 m
2	width	l=0.1 m
3	thickness	t_{fin}=0.002 m
4	numbers of fins	8

Table .2 Heat pipe specification

1	Length of evaporator zone	L_e = 700mm
2	Length of adiabatic zone	L_a = 300mm
3	Length of condenser zone	L_c=1500mm
4	Inner diameter of the pipe	di= 26.6mm
5	Outer diameter of the pipe	do = 33mm
6	Thermal conductivity of the pipe	kiron = 73
7	Thermal conductivity of the insulator	kins =.0206

Purging and Filling the apparatus:

The heat pipe is filled with water through tube A in **Fig.B** to a given level, say to the top of the evaporator or to (-5) cm from the top. Valves 1and 2 are now open. Closing valve 2 and turning two heaters on, the water inside the evaporator is allowed to boil at atmospheric pressure and the vapor leave the system through valve 1 and will carry

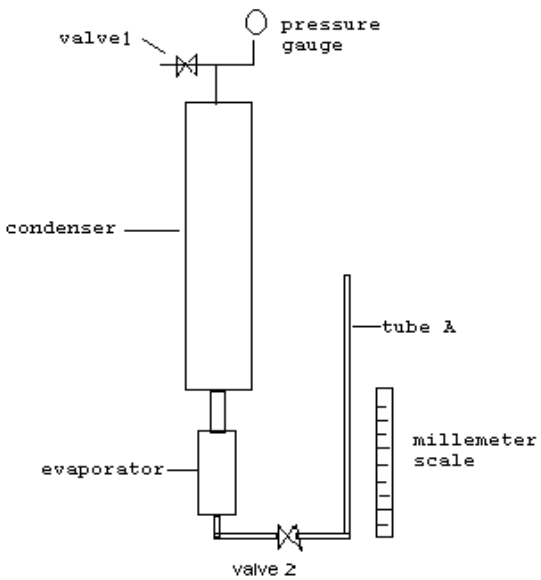


Fig.B: Schematic Diagrams Showing Purging and Filling the H.P with Water

with it the entrapped air. The process of purging the system out of air is continued for one hour. Then valve 2 is opened and distilled water is added to bring the water to the required level then valve 2 is closed. It is to be noted that valve 1 and 2 are of the ball type.

THEORETICAL MODEL

A theoretical steady state model that describes the performance of a heat pipe (H.P) under a variety of operating conditions is developed below. The heat pipe has a diameter of (d_i=0.0266m,d_o=0.033m) and is (2.5 m) long with an evaporator that is driven by a constant heat flux of length (0.7m) and longitudinally finned condenser that reject heat with the environment through natural conviction. The length of the condenser is (1.5m). Between the evaporator and condenser; there is an adiabatic zone (i.e. insulated zone) of length (0.3m). Full description of the heat pipe system with the instrumentations was given earlier. The variable operating conditions include different input heat flux, fluid charge (water) and inclination of the heat pipe with the vertical.

The net heat input (Q) to the evaporator is given by:

$$Q = Q_t - Q_l \qquad (1)$$

Where Q_t is given by:

$$Q_t = V*I \qquad (2)$$

$$Q_l = \frac{2\pi K_{ins.} \; Le}{\ln(ro/ri)} \Delta t_{ins}$$

In steady state operation, Q is balanced exactly by the heat rejected by the condenser Q_c. Thus:

$$Q = Q_c = VI - \frac{2\pi \, Kins. \; Le}{\ln(ro/ri)} \Delta tins \qquad (3)$$

The energy transfer from condenser to ambient is given by:

$$Qc = h_c A_c (t_c - t_\infty) \qquad (4)$$

Where Qc = heat flow through the condenser and Ac= the effective surface area of the finned condenser $A_c = \pi d_o L_c E \eta_s$. Here **E** is the area enhancement factor, **(Dunn and Reay 1976)**:

$$E = \frac{\text{Area of condenser surface with fins}}{\text{Area of condenser surface without fins}}$$

and η_s is the surface efficiency

$$\eta_s = 1 - \frac{A_f}{A_{oc}} (1 - \eta_F) \qquad (5)$$

Where η_f is the fins efficiency

$$\eta_f = \frac{\tanh(mL)}{}$$

(mL)

Where:

$$mL = l^{3/2}\sqrt{\frac{2\,h_o^*}{K_f\, l\, t_{fin}}}$$

The heat rejected by the condenser to the ambient, may by written as:

$$Q_c = A_c\, h_c\, (t_c - t_\infty) + A_c\, \sigma\, (t_c^4 - t_\infty^4) \qquad (6)$$

Eq. (5) may be put in a more convenient form by defining an effective heat transfer coefficient h_o^* given by **(Zeina 2005)**:

$$h_o^* = (\, h_c + \sigma(\, t_c^3 + t_c^2\, t_\infty + t_c\, t_\infty^2 + t_\infty^3\,))$$

Thus **eq. (5)** may read now:

$$Q_c = A_c\, h_c^*\, (t_c - t_\infty) \qquad (7)$$

And **eq.(3)** may simplify to:

$$A_c\, h_c^*(t_c - t_\infty) = VI - \frac{2\pi\, K_{ins}.\ \ L_e}{\ln(r_o/r_i)}\, \Delta t_{ins} \qquad (8)$$

The natural convection heat transfer coefficient hc may be calculated from the following correlation, El-Wakil M. M., (1988), which is suitable be for constant heat flux and for the range $2*10^{13} < Gr^* Pr < 10^{16}$, Incorporeal, F.P., and Dewitt D.P., (1990).

$$Nu = \frac{h_c\, L_c}{k} = (0.17)\,(Gr^* pr)^{1/4} \qquad (9)$$

All physical properties for air and are evaluated at film temperature = $(t_c + t_\infty)/2$

The pressure **p**, of the heat pipe may now be evaluated by back calculating t_g (refer to **Fig.C**)and using the steam tables to find $p=p(t_g)$.

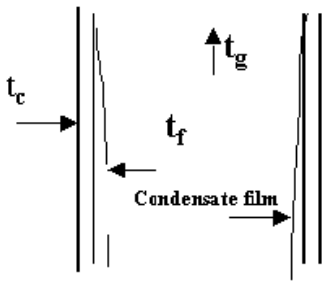


Fig.C: Condenser section showing the definitions of t_c , t_f and t_g.

Using the concept of thermal resistance across the film condensation ,condensate film and wall of the H.P, one easily evaluates t_g as:

$$t_g = t_f + \frac{Q_c}{\ } \qquad (10)$$

$A_i\ h_1$

Where:

$$t_f = t_c + (Q_c / A_i)\left[\left(\frac{r_i}{k_s}\ln\frac{r_o}{r_i} + \frac{r_i}{k_w}\ln\frac{r_i+\delta_w}{r_i}\right)\right] \quad (11)$$

δ_w= average condensate film thickness, **(Incorporeal and Dewitt 1990).**

$$\delta_w = \frac{4}{5}\left[\frac{4\nu_w\ k\ (t_g - t_f)}{g\ Lc^4\ h_{fg}\ (\rho_w - \rho_g)}\right]^{1/4} Lc^{5/4}$$

h_1=the average film condensation heat transfer coefficient and is given by **(Holman 1997):**

$$h_1 = 0.943\left[\frac{k_f^3\ g\ h_{fg}\ (\rho_w - \rho_g)}{\mu_w\ Lc\ (t_g - t_i)}\right]^{1/4}$$

All physical properties for condensate (μ_w, ρ_w, ν_w, h_{fg}) are evaluated at the average condensate temperature (ti+t_g) / 2

Where t_i may readily be written as:

$$t_i = t_c + (Q_c / A_i)\left(\frac{r_i}{k_s}\ln\frac{r_i+\delta_s}{r_i}\right)$$

Equation (9) may now be solved iteratively to determine t_g and hence the pressure of the H.P.

The performance of the H.P may be conveniently characterized by the following efficiency η and thermal resistance, R of the H.P:

$$\eta = \frac{Q_c}{Q_t} \quad (12)$$

$$R = \frac{t_e - t_c}{q_c''} = \frac{(t_e - t_c)}{Q_c}\ Ai \quad (13)$$

RESULTS AND DISCUSSIONS

Experimental results that relates the basic parameters of the H.P (average temperature of condensers surface, average temperature of the evaporator surface and system pressure) under various operating conditions (inventory, net power input to the evaporator and angle of inclination of the H.P with the vertical) are presented and discussed along with performance indices represented by the thermal resistance of the heat pipe and its efficiency. A comparison of both the experimental pressure and condensers temperature with theoretical predictions were presented and discussed.

Transient Behavior of the Heat Pipe:

Fig.1 and **Fig.2** show the transient behavior of a typical run for the evaporator and condenser respectively. The familiar rise to the eventual steady state temperatures is shown. The rise time is around 40-50 minutes. Thus in 2-2.5 hours, the steady state of the system is reached, and hence the steady state measurements may be recorded.

Steady State Behavior of the H.P:

According to the transient time of the H.P discussed previously, all steady state temperatures and pressures were recorded after two hours in half-hour interval and steady state is assume to be reached when successive readings were within $1c^{\circ}$.

Heat Pipe Behavior:

The steady state behavior of the H.P components were taken and include the temperature profiles of the condenser, adiabatic zone and the evaporator respectively. For each component, the average value was computed. The behavior of the H.P was investigated experimentally in relation to: water inventory, net power supplied to the evaporator and angle of inclination of the H.P with the vertical. The results and relevant discussion are given below:

Heat Pipe Behavior with Water Inventory:

When the evaporator is filled to the top, the inventory is said to be 100%. **Fig.3** shows the average temperature of the evaporator (t_e) in relation to percent inventory for different net powers. It is noted that (t_e) is approximately constant (the maximum scatter in the experimental points is ≤ 8% and the constancy is within ± 4%. This behavior is to be expected since, every thing being equal, as the inventory is reduced, the condensation process continues to be of the film type, while the evaporation process changes from totally nucleate boiling to a mixture of nucleate boiling in the submerged depth plus sub-cooled and nucleate boiling in the non sub-cooled boiling occurs as the condensate film descend through the top of the evaporator at a temperature which is lower than saturation. The change in the boiling regime as inventory is reduced does not change the evaporation process significantly since both sub cooled boiling is just as effective as a heat transfer mechanism as that of nucleate boiling. Thus as inventory is reduced, the mechanisms remain essentially the same and subsequently, the temperature (t_c) remains constant. The scatter of experimental points reflects the fact that each experimental point is the result of a given net power and ambient environmental temperature that is slightly different from the next point. These differences, which constitute an inadequacy of the experimental rig, are indicated on **Fig.3** and are unfortunately unavoidable. **Fig.4** shows the average condenser temperature in relation to inventory for different net power to the evaporator. Again, one notes the constancy of the average condensers temperature as the inventory is changed (the mean standard deviation of the points is ≤ 6%). The deviation from constant (t_c) is with in ≤ 5%). The same explanation given earlier for the constancy of (t_e) holds here.

Fig.5 shows the average temperature for the adiabatic zone as a function of inventory and for different net evaporator's power.

Fig.6 shows system pressure for the H.P as a function of inventory for different net evaporator's power. The pressure of the H.P is directly related to (t_c) and

the heat flux, q_c'' leaving through the condenser, thus for constant (t_c) and q_c'', the H.P pressure should have a given value. **Fig.6** shows that the pressure is not constant with inventory, as one would expect if (t_c) and q_c'' were kept constant. But as explained earlier and as indicated on the captions in **Fig.6** both, the net power q_c and (t_∞) vary slightly from one inventory point to the next. As indicated in chapter four (t_g)(and hence the system pressure) is related to (δ_w) and (t_∞). But (δ_w) is directly related to q_c and therefore the shape of the pressure curves indicated in **Fig.6** reflects the combined variations in q_c and (t_∞) as indicated in captions.

Fig.7 shows the thermal efficiency of transferring power through the H.P as a function of inventory, for different net power to the evaporator. Due to slight variation in main voltage, the input power during these measurements fluctuated by few percent and the ambient temperature change by $\pm 5C^\circ$. Since both of these parameters affect the actual heat loss from the evaporator, and subsequently they affect the surface efficiency as defined in **eq. (10).** In split of this fluctuation, the efficiency is almost constant with inventory at all power levels. The maximum standard deviation is ± 3.8% while the best-fit straight lines have slopes ≤ 0.06.

It is to be noted that the efficiency is very high at all powers but being the highest at 99% for an input power 2.25kW. The efficiency is still very high at 96% for an input power of 3kW.

Behavior with Angle of Inclination:

Before attempting to present and discuss the results of the H.P, it is instructive to describe what happens inside the H.P. As the H.P is tilted away from the vertical, the condenser area available for film condensation decreases while that available to drop wise condensation increases.

Since heat transfer coefficient in drop wise condensation is nearly four times the corresponding film wise condensation. And that there is temperature drop across film condensate, while no such drop exists over the region where drop wise condensation is occurring, then it follows, with every things being equal, that condensation occurs at lower temperature as the angle of inclination is increased. This indeed is the case as shown in **Fig.6.** It is known that the condenser controls the pressure of the H.P and since the dominant heat transfer occurs in the region of drop wise condensation where there is no temperature drop across a film then we expect that as the angle of tilt charges. One sees no charge in the surface temperature in that region and this temperature controls the pressure of the H.P. Thus one would expect that with every things being equal, the pressure of the H.P remains constant with tilt angle. This deduction is proved correct in **Fig.9** where the pressure of the heat pipe is shown to be content.

In the evaporator region the picture is more complicated. As the angle of inclination is increased, boiling changes from nucleate boil in the liquid film region plus in the submerged region at angle zero to nucleate boiling in the wetted region plus film boiling in the non wetted or dry region. As the angle of inclination is further increased (i.e. as the heat pipe approaches the horizontal position) then plug flow boiling dominates the entire region of the evaporator since plug flow takes over. Thus one would expect the rate transfer mechanism to decrease initially, then to recover and start to increase. Consequently, the evaporator's temperature (t_g) is expected to rise and eventually to reach a maximum, then declines. This behavior is observed in the experimental H.P as shown in **Fig.10.**

The efficiency of the H.P as a function of inclination angles is shown in **Fig.11.** The efficiency of transferring power actually slightly improves initially (from 96% at 30°angles) then it declines rapidly to below (97% at 60°angles). Thus the heat pipe is very efficient at an inclination angle between (0-30°).

Behavior with Heat Flux (or Net Power):

Fig.12 shows the thermal resistance for the H.P as a function of the heat flux through the condenser at 100% inventory. It is shown that the value of thermal resistance is reasonably small up to a condenser heat flux of (1.158 kW/m^2) and the resistance then begins to increase sharply with heat flux due to changing the boiling mechanism from nucleate to partial film which is accompanied by sharp increase in (te) and subsequent increase in thermal resistance. Thus a conservative upper limit for normal operation of the heat pipe is at (1.2 kW/m^2).

From **Fig.12** one may consider the last point, which corresponds to heat flux (1.48 kW/m^2) to represent a point very close to the critical flux since further increase in heat flux resulted in heaters burn up. These values of thermal resistances could by drastically reduced (by ten folds for example) if the electric heaters were welded directly to the surface of the boiler instead of being attached, by metal band as in the current experiment. Of course in actual practice gas or kerosene fuels will heat these residential H.P, and then the actual thermal resistance will be many folds lower than obtained in the present experiment.

Fig.13 shows the system pressure (which is solely controlled by the condenser) as a function of heat flux through the condenser at (100%) inventory. The pressure increases with heat flux. We note that the slope of the curve also increases with the heat flux. A reasonable pressure of 1.2 to1.4 Mpa may be chosen as working pressure for a domestic H.P.

In summary, the design heat flux through the condenser should be (≤ 1.158 kW/m^2).

Comparison with the Theoretical Model:

Fig.14 and **Fig.15** depicts the comparison between the actual (tc) measurements and the corresponding theoretical counter part as a function of inventory percent. It is to be noted that the complicated shape traced by the theoretical model is only a reflection of the combined variation in net power (rejected by the condenser) and (t_∞). These variations as mentioned earlier, are beyond the experimenter's control. The maximum standard deviation is (≤ ± 16%). The model tends more often to over estimate the value of (t_c) but non the less the theoretical prediction may by regarded as adequate for design purposes.

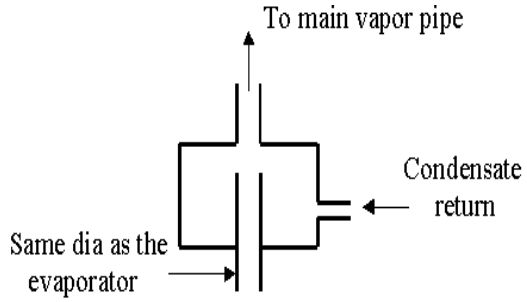
Fig.16 shows the comparison between experimental system pressure and the corresponding theoretical predictions as a function of inventory. It is observed that the theoretical model always underestimate system pressure. Fortunately, the maximum standard deviation is 21% and is not considered very large.

Design of a Residential H.P:

The heat flux of (1.1589 kW/m^2) with the corresponding system pressure of 1500kpa (at 100 % inventory) will be taken as a basis for any design, such as this one. Thus all piping and units should withstand 1.5x design pressure or 2250kpa. No instrumentation is necessary. The closed system should be purged after installation, then filled with distilled water to 50% inventory and then sealed. Actual

implementation of the H.P to heat a residence (composes, say of three spaces with each space requiring 4kW heating) necessitate the following components:

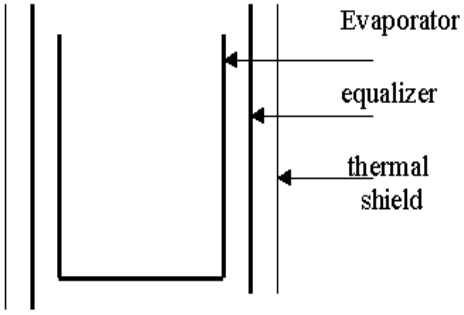
1. *Collector and film flow generator*: This unit which connects to the



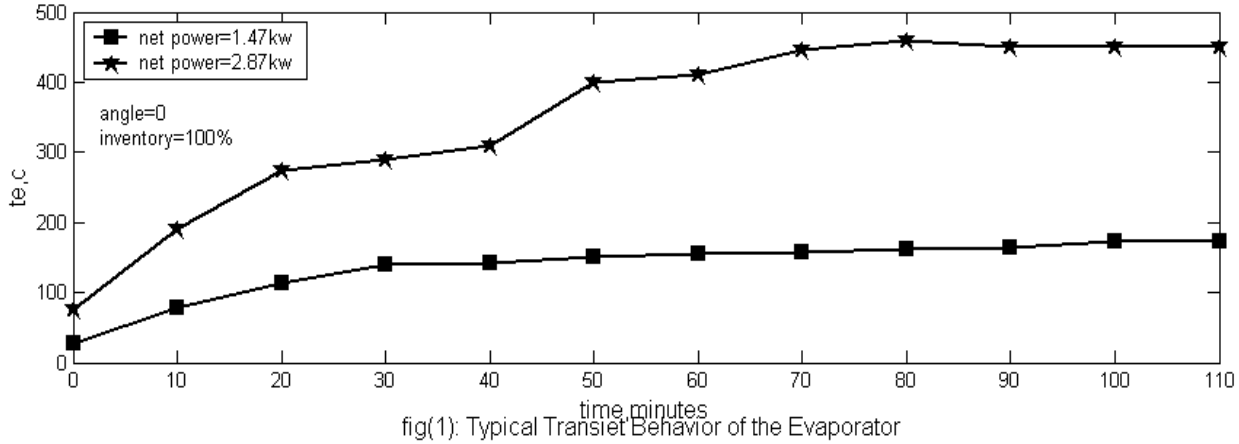
Sketch of the Collector and Film Flow Generator

evaporator simply receives the condensate as a stream and convert it to film flow down the evaporator. This unit should be welded on top of the evaporator and is shown below schematically for one such design.

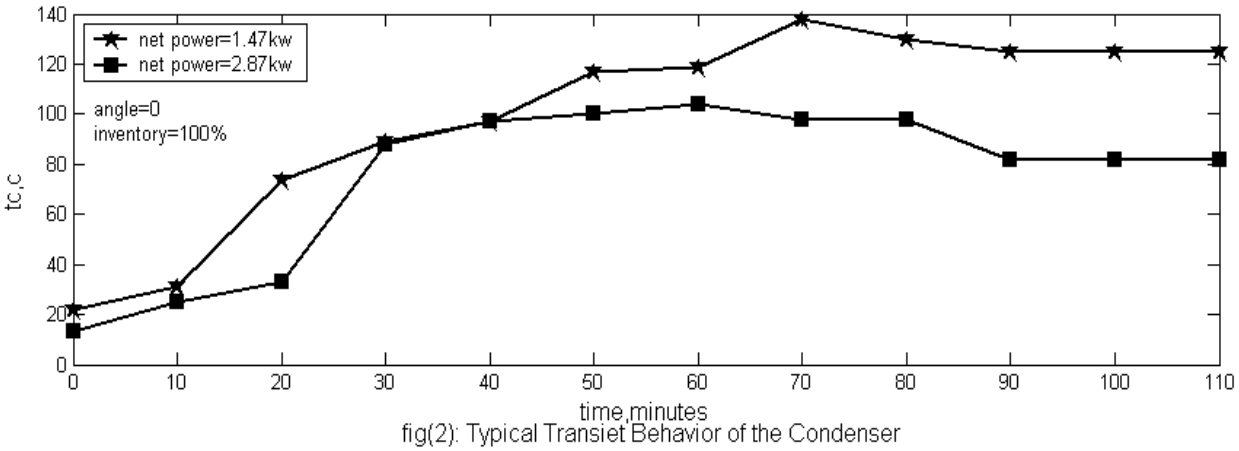
2. *Vapor supply pipe:* This pipe (same diameter as the evaporator pipe) is attached to the top of the collector and supply steam to all the condensers, one for each space. Since in our example we have three condensers, then the branched pipes (pipe to each condenser) have a pipe diameter $1/\sqrt{3}$ =0.577 times the diameter of the main branch. The main branch should be designed for 12 kW and since the experimental H.P (with di of 26.6mm) operated normally up to 1.2 kW/m^2 (i.e. at 3kW total power), then it follows that the main branch should have a di of 26.6 $\sqrt{12/3}$=53.2mm. one important consideration is that this pipe should have a slope that is ≥ 0.2 to facilitate the return of any condensate to the evaporator.
3. *The condenser*: This is a standard finned tube condenser with inlet manifold at the top and condensate out let at the bottom. It is sized on 1.12 kW/m^2. A twenty five percent increase in area is added for conservative design. This gives a total of (4/1.12)*1.25=6 m^2 as the required surface area for each condenser. It should be noted here that all exposed piping inside the residence should be considered part of the condenser unit.
4. *Condense pipe*: The main condensate pipe receives the condensate from the individual condensers. The diameter of the main pipe should be sized to give water velocity of 1m/minute. This requires a diameter of ≥21.7mm. The branch pipes required have a diameter of $21.7\sqrt{1/3}$ mm. All condensate pipes should have a slope ≥ 0.2.
5. *The evaporator*: This is located outside the residence, and has the main branch should be designed for 12kW and since the experimental H.P (with di of 26.6mm) operated normally up to 1.2 kW/ m^2.(i.e. at 3kW total power), then it follows that the main branch should have a di of $26.6\sqrt{12/3}$= 53.2mm. For gas or kerosene heating, a tube of approximate diameter that is greater than the burner should enclose the boiler. This tube acts as a flux equalizer: to distribute the input heat uniformly along the cylindrical evaporator. A thermal shield is required to enclose the flux equalizer in order to reduce heat losses and improve the thermal efficiency. A sketch is shown below.



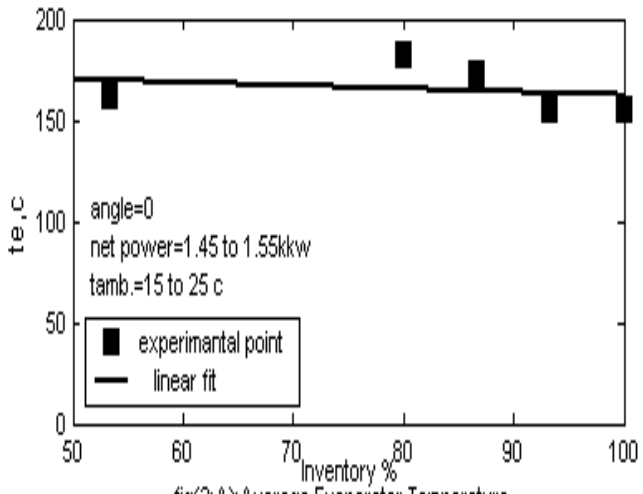
6. The parts of the pipes that are located out side the residence should be insulated.
7. A pressure safety device, such ruptured disk should be attached, very close to the evaporator and is placed out side the building. It is vented to the outdoors. The setting for the ruptured disk should be 1.5 system pressures (~ 1.8mpa).



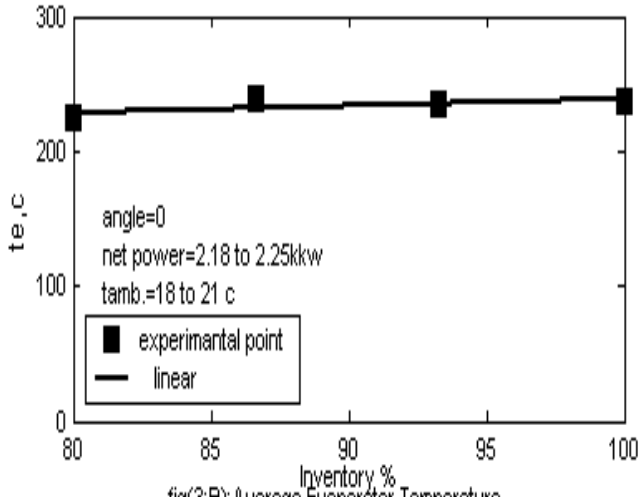
fig(1): Typical Transiet Behavior of the Evaporator



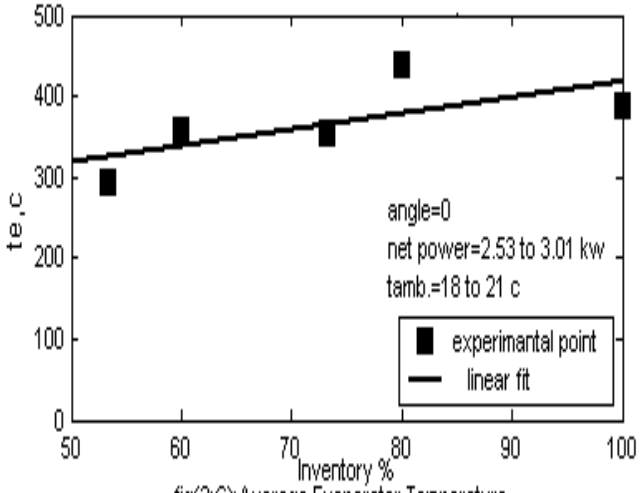
fig(2): Typical Transiet Behavior of the Condenser



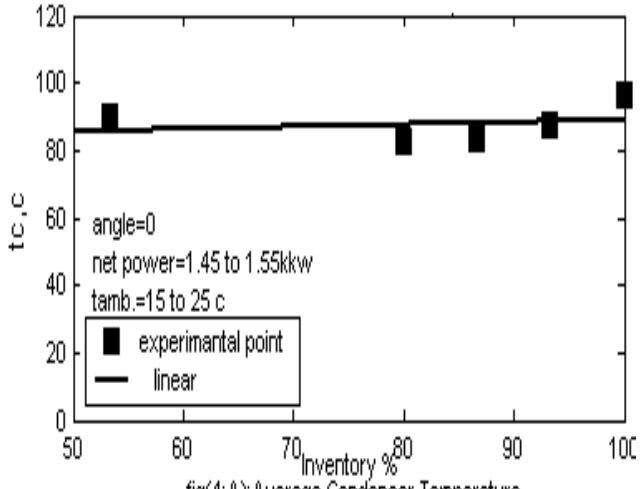
fig(3:A):Average Evaporator Temperature
With Different Water Inventory



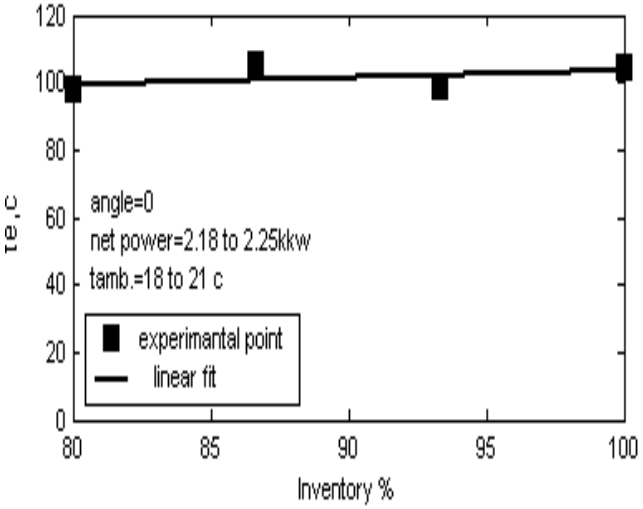
fig(3:B):Average Evaporator Temperature
With Different Water Inventory



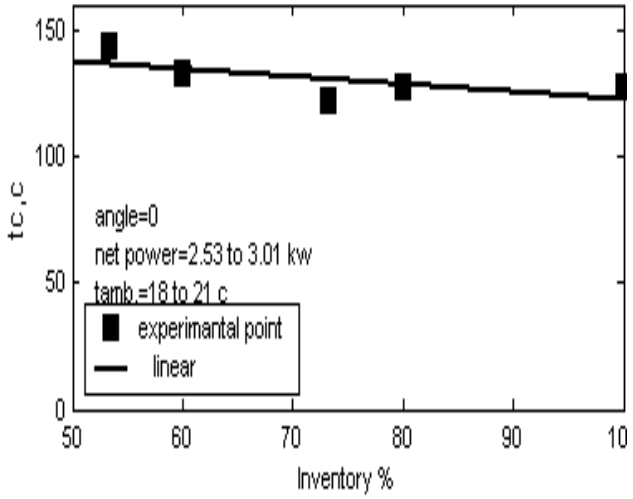
fig(3:C):Average Evaporator Temperature
With Different Water Inventory



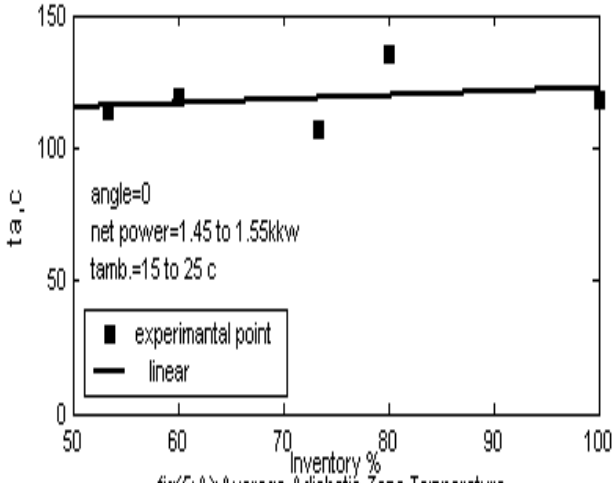
fig(4:A):Average Condenser Temperature
With Different Water Inventory



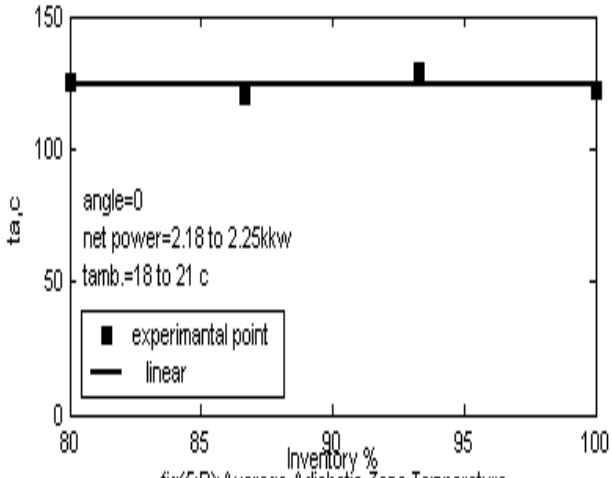
fig($:B):Average Condenser Temperature
With Different Water Inventory



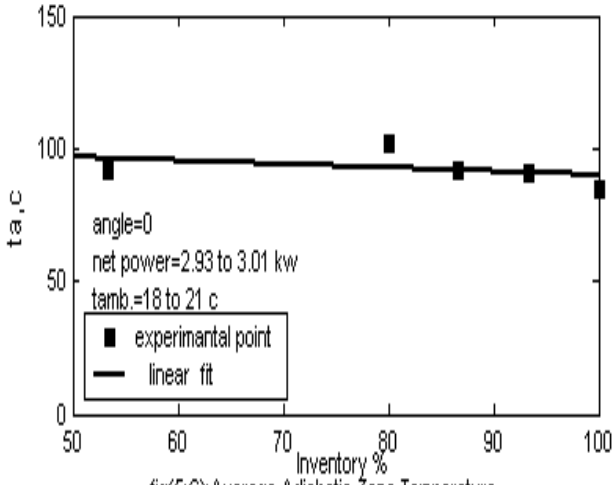
fig(4:C):Average Condenser Temperature
With Different Water Inventory



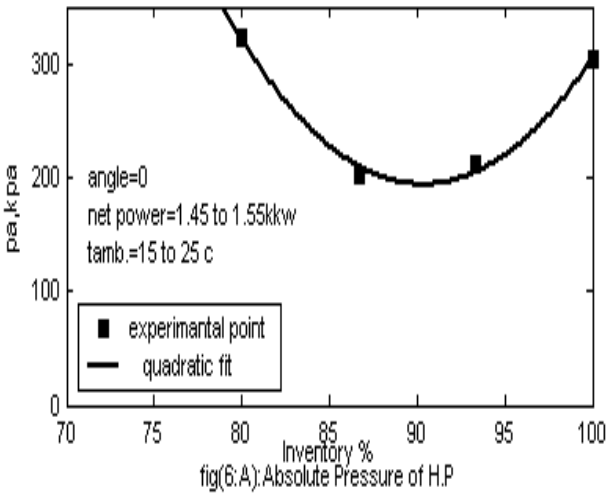
fig(5:A):Average Adiabatic Zone Temperature
With Different Water Inventory



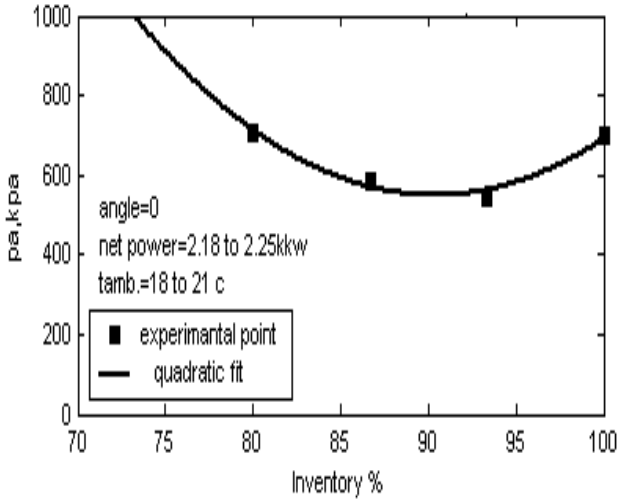
fig(5:B):Average Adiabatic Zone Temperature
With Different Water Inventory



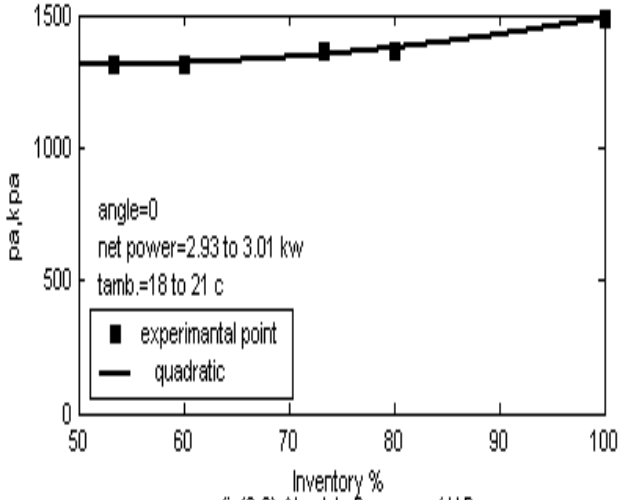
fig(5:C):Average Adiabatic Zone Temperature
With Different Water Inventory



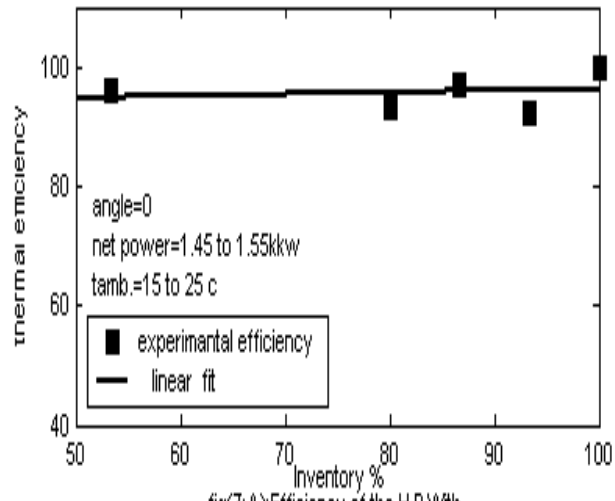
fig(6:A):Absolute Pressure of H.P
With Different Water Inventory



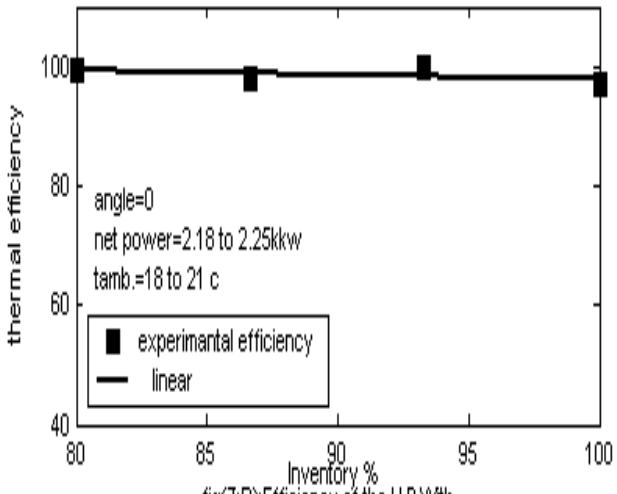
fig(6:B):Absolute Pressure of H.P
With Different Water Inventory



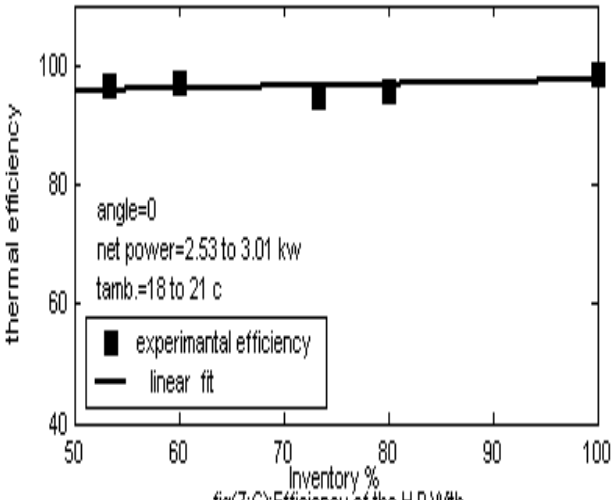
fig(6:C):Absolute Pressure of H.P
With Different Water Inventory



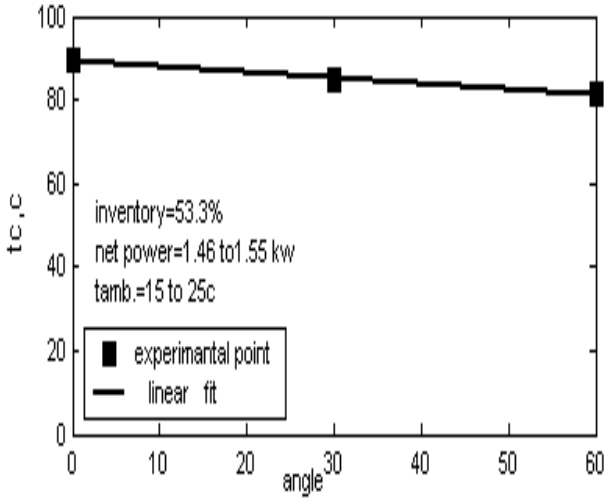
fig(7:A):Efficiency of the H.P With Different Water Inventory



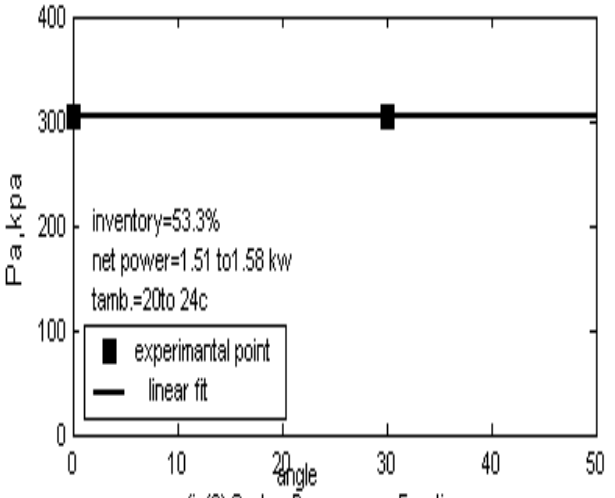
fig(7:B):Efficiency of the H.P With



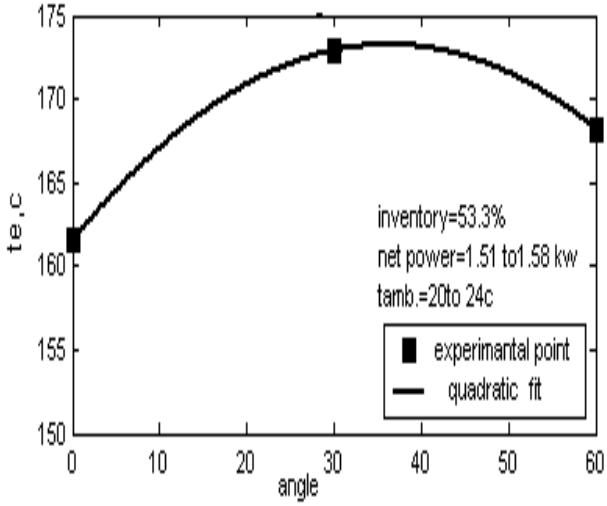
fig(7:C):Efficiency of the H.P With Different Water Inventory



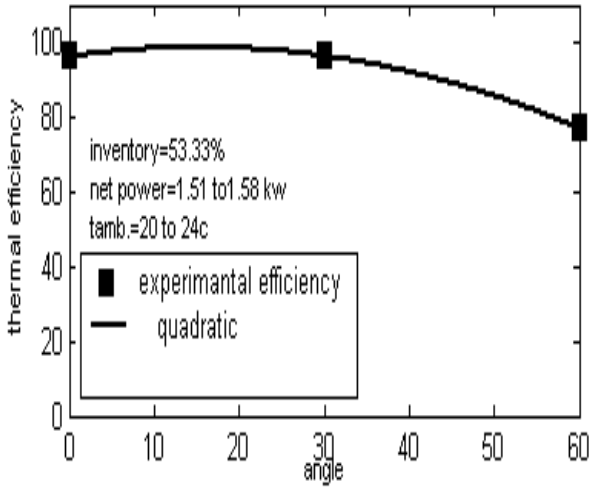
fig(8):Average Condenser Temperature For Different Angle of Inclination



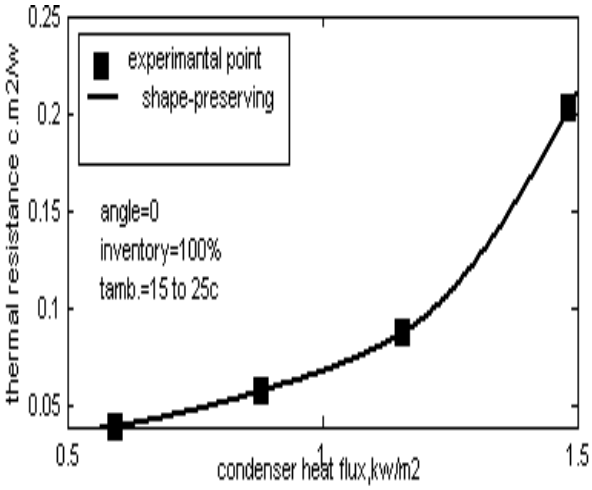
fig(9):System Pressure as aFunction of Inclination Angle



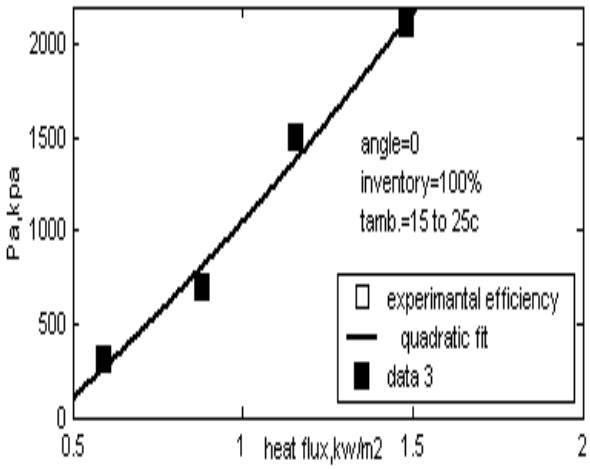
fig(10):Average Evaporator Temperature For Different Angle of Inclination



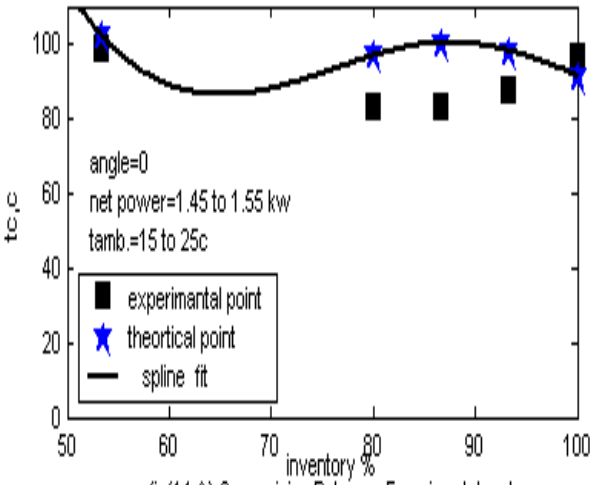
fig(11):Efficiency Relation to Angle of Inclination



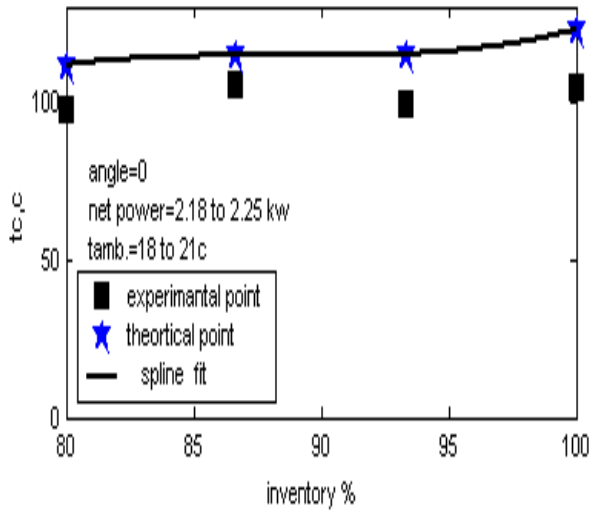
fig(12):Thermal Resistance of the H.P With Condenser Heat Flux



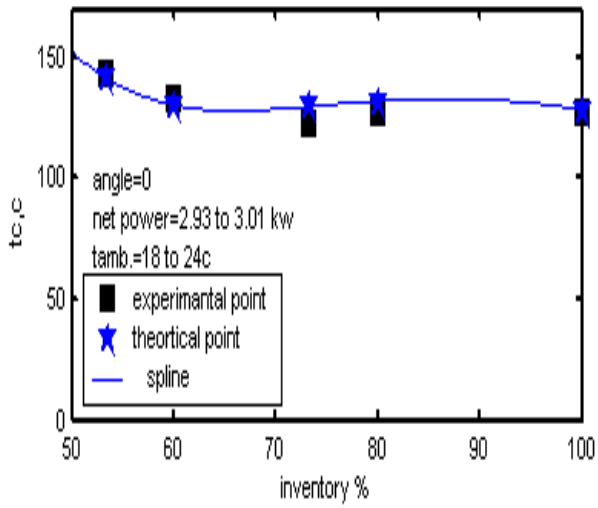
fig(13):Efficiency of the H.P With Different Water Inventory



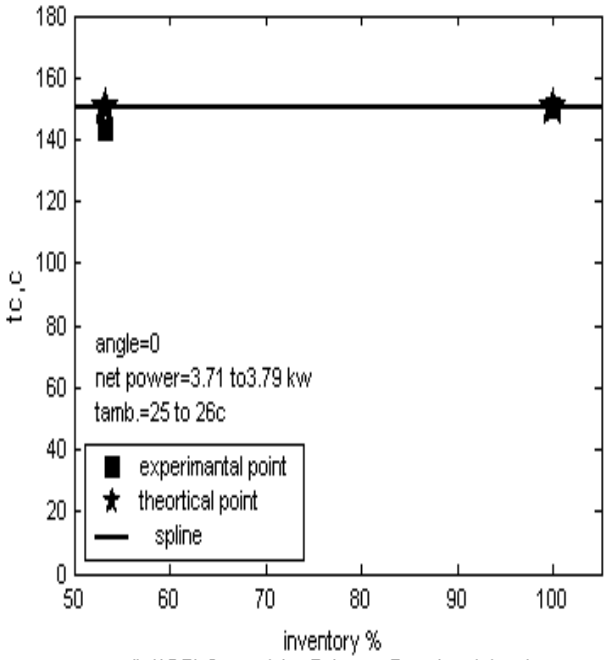
fig(14:A):Comparision Between Experimantal and Theortical Average Condenser Temperature



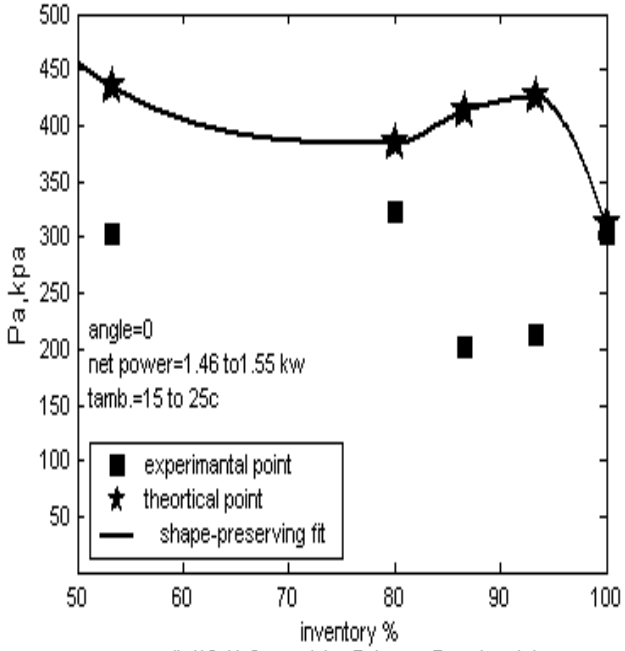
fig(14:B):Comparision Between Experimantal and Theortical Average Condenser Temperature



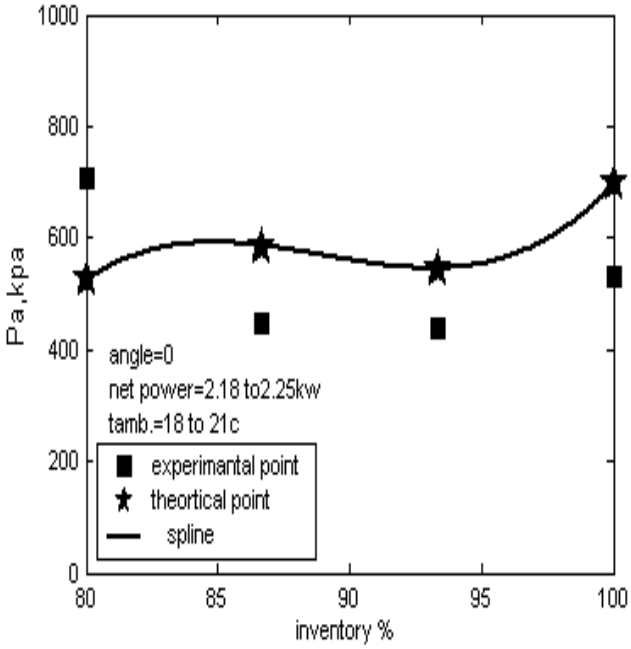
fig(15):Comparision Between Experimantal and Theortical Average Condenser Temperature



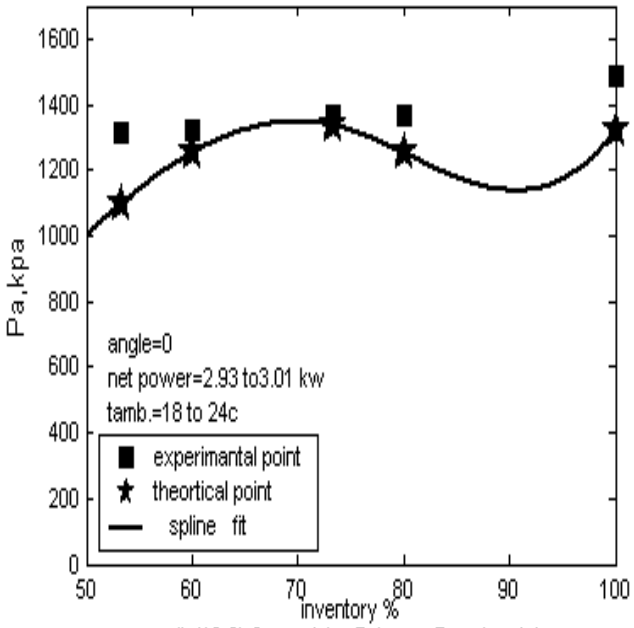
fig(15:B):Comparision Between Experimantal and Theortical Average Condenser Temperature



fig(16:A):Comparision Between Experimantal and Theortical Absolute Pressure



fig(16:B):Comparision Between Experimantal and Theortical Absolute Pressure



fig(16:C):Comparision Between Experimantal and Theortical Absolute Pressure

CONCLUSIONS:

1. For a given (vertically oriented) H.P design, the important operating parameters of (p, t_c ,q_c'') are found to be independent of total inventory. Practical consideration, however, favor an inventory close to 50 %. Since lower inventory may cause unwarranted high evaporator's temperature during start up while higher inventory may reduce the part of the evaporator where boiling film flow takes place.
2. A conservative design flux for the condenser unit (in vertically oriented H.P) is 1.2 kW/m^2. This flux is nearly 30 % lower than the critical flux for the H.P. This design flux, which controls the system pressure, maintains a system pressure of 1500 kpa.
3. For an inclined H.P, the performance (as reflected in efficiency) is slightly improved with angle of inclination equal to 30°. The pressure of the system is totally independent of inclination angle.
4. The theoretical model developed to predict the condensers temperature and system pressure for a given net power input is accurate to ± 16% in predicting t_c and with ±21% in predicting system pressure. This accuracy is adequate for design purposes.

REFERANCES:

Chi, S.W.,**' Heat Pipe Theory and Practice'**, 1st ed., Hemisphere Publishing corporation, Washington-London, 1976.

Dunn, P.D. and D.A. Reay, '**Heat Pipe'**, 1st ed., A.Wheaton and Co.Exeter, Great Britain, 1976.

El-Wakil M. M., ' **Nuclear Power Engineering**', 8 th ed., Mc Graw-Hill Book Company, 1988.

Holman, J.P., **'Heat Transfer'**; 8th ed., McGraw-Hill Kogakusha, 1997.

Incorporeal, F.P. and D.P.Dewitt,**'Fundamentals of Heat and Mass Transfer',** 3d ed., Johan Wiley and Sons, 1990.

Marquet, C. and J.C. Solecki etall,**'Heat transfer Performance of a Toluene Loaded Two-Phase Thermosyphon'**, Heat Recovery Systems and CHP, Vol. 4,PP. 285-297,1989.

Monem, H.B. and C.D. Patel etall,**'Desigm and Performance Evaluation of a Compact Thermosyphom'**, Hewlett-Packard Research Laboratories, California, 2003.

Qian, w. and g. Chen,'**Analyses of Heat Pipe',** ASME J.Terbologe, Vol.121, Pp.551-557, 1999.

Shunji, K. and H. Suzuki etall, **'Extraction of Geothermal Energy and Electric Power Generation Using a long Scale Heat Pipe'**, Proceedings World Geothermal Congress, Kyushu-Tohoku, 2000.

Torrance K.E.,**'Open- Loop Thermosyphons With Geological Application'**, J.Heat T ransfer, vol. 101,pp. 677-683,Nov. 1979.

Zeina, A.AL-Saadi.,' **Design and Implementing a Heat Pipe Experimental System for Residential Heating**', Msc thesis, Nuclear Eng. Dept.,University of Baghdad,2005

NOMENCLATURE:

Symbol	Meaning	Unit
A	area	m^2
C_P	specific heat capacity	kJ/kg.k
d	pipe diameter	m
E	enhancement factor	---------
g	gravitation acceleration=9.8	m/s^2
h	heat transfer coefficient	$W/m^{2 \cdot 0}c$
K	thermal conductivity	$W/m \cdot {}^0c$
l	Fin length	m
L	pipe length	m
Nu	Nusselt number=hd/k	---------
p	power	W
P	pressure	kPa
q"	the heat flux to the evaporator	$kJ/m^2.s$
Q	heat flow	w
r	radius of pipe	m
s	thickness of fins	m
t	temperature	${}^0c$
V	voltage	volt
I	current	ampere
Pr	prandtl number = $\mu C_P /k$	-------
Gr	Grashof number = $g \beta (t_s - t_\infty) L_c^3 / \nu^2$	--------
Gr^*	modified Grashof number = Nu Gr	--------
h_{fg}	Heat of vaporization	J/kg

Greek Symbol:

Symbol	Meaning	Unit
β	Volume expansion coefficient for steam	1/k
μ	dynamic viscosity	$N.s/m^2$
ρ	density	kg/m^3
ν	kinematics viscosity	m^2/s
δ	Stefan-Boltzman constant $=5.669*10^{-8}$	$W/m^2 .k^4$
∞	ambient	--------
θ	angle	degree
Δ	change (t_2-t_1)	---------
η	efficiency	---------
φ	Heat flux	W
δ	average liquid film thickness	---------

Subscripts:

Symbol	Meaning
a	adiabatic zone
c	condenser
e	evaporator
f	fluid
fin	Fins surface area
g	vapor
ins	insulator
i	inner
l	losses

o	outer
s	cast iron
t	total
uf	un finned
w	water
x	local

MACHINE-TOOL SETTINGS TO PROVIDE OPTIMAL TCA OF SPIRAL BEVEL GEAR DRIVES

Mohammad Qasim Abdullah
Assist. Prof. in Mechanical Engineering Department
College of Engineering / University of Baghdad

ABSTRACT

The present paper presents an approach to optimize tooth contact analysis (TCA) of spiral bevel gear drivin by controlling the machine-tool settings that directly affects the shape of tooth and behavior of meshing and contact for mating gears. The proposed settings provide a pre-designed parabolic function of transmission errors and the desired location and orientation of the bearing contact. The main goal of detecting the pre-designed parabolic function of transmission errors is to reduce the gear noise which can be done by absorbing the linear function of transmission errors that are caused by gear misalignment. The model is generated with means of CAD software package and solid works program, the basic input design data imported by Gleason works standards.

الخلاصة:

في هذا البحث تم تقديم وسيلة لإيجاد أفضل تحليل لتلامس الأسنان للمسنن المخروطي الحلزوني وذلك بواسطة التحكم بضبط موقع الأداة لماكنة القطع الخاصة بهذا النوع من المسننات والتي تؤثر بشكل مباشر على شكل السن وسلوك المحاكاة والتلامس بين المسننات المتعاشقة. إن الضبط المقترح لأداة القطع سوف يأخذ بنظر الاعتبار الأخطاء المتولدة من جراء آلية نقل الحركة وكذلك موقع التحميل ومنحى التلامس الحاصل بين الأسنان. إن الهدف الرئيسي من هذا البحث هو تقليل الخطأ الحاصل من جراء آلية نقل الحركة والمتسبب من عدم التطابق بين المسننات المتعاشقة، وذلك لتخفيض مستوى الضوضاء المتولدة منها عند التعشيق. ولغرض الشروع في هذه الدراسة، تم استخدام برنامج جاهز لتوليد النماذج الخاصة بهذا النوع من المسننات وفقاً لبيانات إدخال قياسية استخدمت لهذا الغرض.

KEYWORDS: Gear, Spiral, Bevel, Tool, Setting, TCA.

INTRODUCTION

Tooth Contact Analysis (TCA) is a computational approach for analyzing the nature and quality of the meshing contact in a pair of gears. The concept of TCA was originally introduced in early 1960s as a research tool, and applied to spiral bevel gears. It is a powerful tool for the design and analysis of spiral bevel gear drives. Typical outputs of TCA are the graphs of contact patterns and transmission errors.

TCA can simulate gear meshing contact characteristics under light and heavy loads. TCA program have been widely employed by gear engineers and researchers in their design of high strength and low noise spiral bevel gear drives.

Application of TCA technology resulted in significant improvement in the development of bevel gear pairs, under given contact conditions **(Lelkes and Marialigeti , 2002).**

Basically, machine tool setting means the guide to design and manufacture the gear drive. In the present work, a developed approach has been proposed to control and design suc a gear drive by simulating the meshing and changing the machine-tool settings to get optimal TCA.

There are two methods for manufacturing spiral bevel gears, face milling and face hobbing. Both of which are widely employed by the gear manufacturing industry and can be implemented on modern CNC bevel gear generators **(Litvin and Lee , 1989).**

Fig.(1) shows a 3D geometric model of spiral bevel gear.

MACHINE-TOOL SETTENGS AND TCA

Gear Machine-Tool Settings

The proposed design provides the following:

The gear-generating surface by $\sum_G$, the generated gear surface by $\sum_2$, the pinion-generating tool surface by $\sum_P$, and the generated pinion surface $\sum_1$.

To set up the gear machine-tool settings, the following data should be given:

Γ: shaft angle
N_2: gear tooth number, N_1: pinion tooth number
γ_2: gear root angle
A: mean pitch cone distance
β: mean spiral angle
ψ_G: blade angle for gear cutter
d_G: average diameter of gear cutter
w_G: point width

The gear pitch angle is represented by (Faydor L. Litvin and Alfonso Fuentes, 2004):

$$\mu_2 = \arctan \frac{\sin \Gamma}{\frac{N_1}{N_2} + \cos \Gamma} \quad (1)$$

The pinion pitch angle is:

$$\mu_1 = \Gamma - \mu_2 \quad (2)$$

The dedendum angles are:

$$\delta_1 = \mu_1 - \gamma_1 \text{ and } \delta_2 = \mu_2 - \gamma_2 \quad (3)$$

Gear Cutting Ratio

The process of gear generation is based on the imaginary meshing of a crown-gear with the member-gear **(Qi Fan 2006).**

The instantaneous axis of rotation by such meshing coincides with the pitch line axis Z_P, as shown in **Figs. (3)** and **(4).**

The generating surface $\sum_G$ which may be imagined as the surface of the crown gear, and be generated gear surface $\sum_2$ contact each other at a line at every instant. The ratio of angular velocities of the crown gear and the being generated gear (the cutting ratio) remains constant while the spatial line of contact moves over surfaces $\sum_G$ and $\sum_2$.

The gear cutting ratio can be represented as follows **(Litvin and Lee 1989):**

$$m_{G2} = \frac{\omega^{(G)}}{\omega^{(2)}} = \frac{\sin \mu_2}{\cos \delta_2} \quad (4)$$

Cutter Tip Radius, Radial Setting, and Cradle Angle

From **Fig.(5)**, it can be obtained the inside and outside tip radii of the head-cutter as follows **(Litvin and Lee 1989):**

$$r_G = \frac{1}{2}(d_G \mp w_G) \quad (5)$$

Also, from the relation between the lengths and angles of the triangle $O_m O_c M_G$ of the same figure, it can be expressed the radial setting S_G and cradle angle q_G as follows:

$$S_G = \sqrt{\frac{d^2}{4} + A^2 \cos^2 \delta_2 - d_G A \cos \delta_2 \sin \beta} \quad (6)$$

and

$$q_G = \arccos \frac{A^2 \cos^2 \delta_2 + S_G^2 - d_G^2 / 4}{2 A S_G \cos \delta_2} \quad (7)$$

DETERMINATION ANALYSIS OF THE MEAN CONTACT POINT

The gear and pinion surfaces of spiral bevel gears are in point contact at every instant. The mean contact point is the center of the bearing contact and its location is selected generally at the middle of the working depth on gear tooth. **Fig.(6)** shows a gear tooth surface. Section AD is the gear tip, section BC is the pinion tip and it is parallel to the root line of the gear, and the working area is within ABCD **(Faydor and Alfonso 2003).**

The mean contact point is located on a line which passes through the middle point of the two points at which the normal section of the gear surface intersects line AD and line BC respectively. In addition, the mean contact point must be on the gear surface. This means that is must satisfy the equation of meshing for the gear being generated by the tool. From these two requirements, the location of the mean contact point is determined **(Joseph and Thomas 2003).**

RELATION BEWEEN DIRECTIONS OF THE PATHS OF THE MEAN CONTACT POINT OVER THE GEAR AND PINION TOOTH SURFACES

Fig.(7) shows the tangent plane to the gear and pinion surface at the mean contact point B. The relation between angles υ_1 and υ_2 depends on parameters in motion and the principal curvatures of the gear tooth surface.

This relation can be expressed as follow (Gosselin C, Cloutier L and Brousseau J., 1991):

$$\tan\upsilon_1 = \frac{(a_{33} + a_{31}V_{2I}^{(12)})\tan\upsilon_2 - a_{31}V_{2II}^{(12)}}{a_{32}(V_{2II}^{(12)} - V_{2I}^{(12)}\tan\upsilon_2) + a_{33}} \tag{8}$$

PINION MACHINE – TOOL SETTINGS

There are five machine-tool settings m_{p1}, E_m, L_m, s_p, and q_p to be determined. The key to the solution of this problem is the determination of the cutting ratio m_{p1} (Lelkes M. and Marialigeti J., 2002).

Determination of Pinion Cutting Ratio

Consider that surfaces $\sum_1$ and $\sum_F$ are equivalent, and that surfaces $\sum_p$ and $\sum_Q$ are equivalent.

Also the following data must be given **(Lelkes and Marialigeti 2002):**

1- The principal curvatures of the pinion surface at the mean contact point, k_{1I} and k_{1II} .
2- The principal directions of the pinion surface at the mean contact point, $\vec{e}_{1I}$ and $\vec{e}_{1II}$.
3- The coordinates of the mean contact point.
4- The unit normal at the mean contact point.
5- The coefficients a_{11}, a_{12}, and a_{22}.

The procedure to determine m_{p1} is as follows **(Robert Handschuh 1997):**

<u>Step 1:</u> The angular velocity of the pinion is represented by

$$\vec{\omega}^{(1)} = \pm \begin{bmatrix} \sin\mu_1 \\ 0 \\ \cos\mu_1 \end{bmatrix} \tag{9}$$

The angular velocity of the pinion cutter is represented by:

$$\vec{\omega}^{(p)} = \pm m_{p1}\omega^{(1)} \begin{bmatrix} \cos\delta_1 \\ 0 \\ -\sin\delta_1 \end{bmatrix} \tag{10}$$

Therefore, the relative angular velocity $\vec{\omega}^{(1p)}$ can be obtained as follows:

$$\vec{\omega}^{(1p)} = \pm\omega^{(1)} \begin{bmatrix} \sin\mu_1 - m_{p1}\cos\delta_1 \\ 0 \\ \cos\mu_1 + m_{p1}\sin\delta_1 \end{bmatrix} \tag{11}$$

<u>Step 2:</u> Representation of $\left[\vec{\omega}^{(1p)}\vec{n}\vec{e}_{pI}\right]$

The scalar $\left[\vec{\omega}^{(1p)}\vec{n}\vec{e}_{pI}\right]$ is represented by:

$$\left[\vec{\omega}^{(1p)}\vec{n}\vec{e}_{pI}\right] = \omega^{(1)} \begin{vmatrix} \pm(\sin\mu_1 - m_{p1}\cos\delta_1) & 0 & \pm(\cos\mu_1 + m_{p1}\sin\delta_1) \\ n_x & n_y & n_z \\ e_{PIx} & e_{PIy} & e_{PIz} \end{vmatrix}$$

$$= \begin{Bmatrix} \left[(n_z e_{PIy} - n_y e_{PIz})\cos\delta_1 + (n_x e_{PIy} - n_y e_{PIx})\sin\delta_1\right] m_{p1} \\ + \left[(n_y e_{PIx} - n_z e_{PIy})\sin\mu_1 + (n_x e_{PIy} - n_y e_{PIx})\cos\mu_1\right] \end{Bmatrix} \omega^{(1)}$$

$$= (C_{11} m_{p1} + C_{12})\omega^{(1)} \tag{12}$$

<u>Step 3</u>: Representation of $\left[\vec{\omega}^{(1p)}\vec{n}\vec{e}_{PII}\right]$

The scalar $\left[\vec{\omega}^{(1p)}\vec{n}\vec{e}_{PII}\right]$ is represented by:

$$\left[\vec{\omega}^{(1p)}\vec{n}\vec{e}_{PII}\right] = \omega^{(1)} \begin{vmatrix} \pm(\sin\mu_1 - m_{p1}\cos\delta_1) & 0 & \pm(\cos\mu_1 + m_{p1}\sin\delta_1) \\ n_x & n_y & n_z \\ e_{PIIx} & e_{PIIy} & e_{PIIz} \end{vmatrix}$$

$$= \pm\left[(n_y e_{PIIz} - n_z e_{PIIy})\sin\mu_1 + (n_x e_{PIIy} - n_y e_{PIIx})\cos\mu_1\right]\omega^{(1)}$$

$$= C_{22}\omega^{(1)} \tag{13}$$

<u>Step 4</u>: Representation of $\vec{V}^{(1p)}$

The velocity $\vec{V}^{(1)}$ may be obtained by:

$$\vec{V}^{(1)} = \vec{\omega}^{(1)} x \vec{B}$$

$$= \pm\omega^{(1)} \begin{bmatrix} -B_y \cos\mu_1 \\ B_x \cos\mu_1 - B_Z \sin\mu_1 \\ B_y \sin\mu_1 \end{bmatrix} \tag{14}$$

The velocity $\vec{V}^{(p)}$ may be obtained by:

$$\vec{V}^{(p)} = \vec{\omega}^{(p)} x \vec{B} + \overline{O_f O_m} x \vec{\omega}^{(p)}$$

$$= \omega^{(1)} m_{p1} \begin{bmatrix} (E_m \pm B_y)\sin\delta_1 \\ \pm(L_m - B_x \sin\delta_1 - B_y \cos\delta_1 \\ (E_m \pm B_y)\cos\delta_1 \end{bmatrix} \tag{15}$$

So, the sliding velocity $\vec{V}^{(1p)}$ is described by:

$$\vec{V}^{(1p)} = \vec{V}^{(1)} - \vec{V}^{(p)}$$

$$= \omega^{(1)} \begin{bmatrix} \pm B_y \cos\mu_1 - m_{p1}(E_m \pm B_y)\sin\delta_1 \\ \pm(B_x \cos\mu_1 - B_z \sin\mu_1) \pm m_{p1}[L_m - (B_x \sin\delta_1 + B_z \cos\delta_1)] \\ \pm B_y \sin\mu_1 - m_{p1}(E_m \pm B_y)\cos\delta_1 \end{bmatrix} \tag{16}$$

<u>Step 5</u>: Representation of $V_{PI}^{(1p)}$ and $V_{PII}^{(1p)}$

$$a_{13} = -k_{PI} V_{PI}^{(1p)} - (C_{11} m_{p1} + C_{12})\omega^{(1)} \tag{17}$$

$$a_{23} = -k_{PII} V_{PII}^{(1p)} - C_{22}\omega^{(1)} \tag{18}$$

$$a_{11} a_{23} - a_{12} a_{13} = 0 \tag{19}$$

By using **Eqs. (17), (18),** and **(19):**

$$a_{12} k_{PI} V_{PI}^{(1p)} - a_{11} k_{PII} V_{PII}^{(1p)} = \left[-a_{12} C_{11} m_{p1} + (a_{11} C_{22} - a_{12} C_{12})\right]\omega^{(1)} \tag{20}$$

Also, it can be deduced that:

$$\vec{V}^{(1p)} = V_{PI}^{(1p)} \vec{e}_{PI} + V_{PII}^{(1p)} \vec{e}_{PII} \tag{21}$$

By considering only the x component in **Eqs. (16)** and **(21):**

$$V_{PI}^{(1p)} e_{PIx} + V_{PII}^{(1p)} e_{PIIx} = [\pm B_y \cos\mu_1 - m_{p1}(E_m \pm B_y)\sin\delta_1]\omega^{(1)} \tag{22}$$

Also by considering only the z component in **Eqs. (16)** and **(21):**

$$V_{PI}^{(1p)} e_{PIz} + V_{PII}^{(1p)} e_{PIIz} = [\pm B_y \sin\mu_1 - m_{p1}(E_m \pm B_y)\cos\delta_1]\omega^{(1)} \tag{23}$$

By multiplying **Eq. (22)** by cosδ1 and **Eq. (23)** by sinδ1, and adding the resulting equations:

$$V_{PII}^{(1p)} = \pm \frac{B_y \cos\gamma_1}{e_{PIIx}\cos\delta_1 - e_{PIIz}\sin\delta_1}\omega^{(1)} = t_1\omega^{(1)} \tag{24}$$

Then, substituting **Eq.(24)** into **Eq.(22):**

$$V_{PI}^{(1p)} = (-\frac{C_{11}}{k_{PI}}m_{p1} + \frac{a_{11}k_{PII}t_1 + a_{11}C_{22} - a_{12}C_{12}}{a_{12}k_{PI}})\omega^{(1)}$$

$$= (t_1 m_{p1} + t_2)\omega^{(1)} \tag{25}$$

<u>Step 6:</u> Representation of $\vec{V}^{(1p)}$

The matrix form of **Eq.(21)** may be represented by:

$$\vec{V}^{(1p)} = \begin{bmatrix} V_{PI}^{(1p)}e_{PIx} + V_{PII}^{(1p)}e_{PIIx} \\ V_{PI}^{(1p)}e_{PIy} + V_{PII}^{(1p)}e_{PIIy} \\ V_{PI}^{(1p)}e_{PIz} + V_{PII}^{(1p)}e_{PIIz} \end{bmatrix} \tag{26}$$

By substituting **Eqs. (24)** and **(25)** into **Eq.(26):**

$$\vec{V}^{(1p)} = \omega^{(1)}\begin{bmatrix} (t_1 m_{p1} + t_2)e_{PIx} + t_4 e_{PIIx} \\ (t_1 m_{p1} + t_2)e_{PIy} + t_4 e_{PIIy} \\ (t_1 m_{p1} + t_2)e_{PIz} + t_4 e_{PIIz} \end{bmatrix}$$

$$= \omega^{(1)}\begin{bmatrix} u_{11}m_{p1} + u_{12} \\ u_{21}m_{p1} + u_{22} \\ u_{31}m_{p1} + u_{32} \end{bmatrix} \tag{27}$$

<u>Step 7:</u> Representation of $\left[\vec{n}\vec{\omega}^{(1p)}\vec{V}^{(1p)}\right]$

The scalar $\left[\vec{n}\vec{\omega}^{(1p)}\vec{V}^{(1p)}\right]$ may by represented by:

$$\left[\vec{n}\vec{\omega}^{(1p)}\vec{V}^{(1p)}\right] = [\omega^{(1)}]^2 \begin{vmatrix} n_x & n_y & n_z \\ \pm(\sin\mu_1 + m_{p1}\cos\delta_1) & 0 & \pm(\cos\mu_1 + m_{p1}\sin\delta_1) \\ u_{11}m_{p1} + u_{12} & u_{21}m_{p1} + u_{22} & u_{31}m_{p1} + u_{32} \end{vmatrix}$$

$$= (\upsilon_1 m_{p1} + \upsilon_2 m_{p1} + \upsilon_3)[\omega^{(1)}]^2 \tag{28}$$

Where:

$$\upsilon_1 = \pm\left[(u_{11}\sin\delta_1 - u_{31}\cos\delta_1)n_y - (n_z\cos\delta_1 + n_x\sin\delta_1)u_{21}\right] \tag{29}$$

$$\upsilon_2 = \mp\begin{bmatrix} (u_{21}\cos\mu_1 + u_{22}\sin\delta_1)n_x - (u_{21}\sin\mu_1 - u_{22}\cos\delta_1)n_z \\ -(u_{11}\cos\mu_1 + u_{12}\sin\delta_1 + u_{32}\cos\delta_1 - u_{31}\sin\mu_1)n_y \end{bmatrix} \tag{30}$$

$$\upsilon_3 = \mp\left[(u_{22}n_x \cos\mu_1) - (u_{12}\cos\mu_1 - u_{32}\sin\mu_1)n_y - u_{22}n_z \sin\mu_1\right] \tag{31}$$

Step 8: Representation of $\vec{n}.(\vec{\omega}^{(1)} x \vec{V}^{(p)} - \vec{\omega}^{(p)} x \vec{V}^{(1)})$

The velocity $\vec{V}^{(p)}$ may be described by:

$$\vec{V}^{(p)} = \vec{V}^{(1)} - \vec{V}^{(1p)} \tag{32}$$

Now, substituting **Eqs.(14)** and **(27)** into **Eq.(32):**

$$\vec{V}^{(p)} = \omega^{(1)} \begin{bmatrix} -u_{11}m_{p1} \mp B_y \cos\mu_1 - u_{12} \\ -u_{21}m_{p1} \mp B_z \sin\mu_1 \pm B_x \cos\mu_1 - u_{22} \\ -u_{31}m_{p1} \pm B_y \sin\mu_1 - u_{32} \end{bmatrix} \tag{33}$$

Vector $(\vec{\omega}^{(1)} x \vec{V}^{(p)})$ is represented by:

$$\vec{\omega}^{(1)} x \vec{V}^{(p)} = [\omega^{(1)}]^2 \begin{bmatrix} \pm[u_{21}m_{p1} - (B_z \sin\mu_1 - B_x \cos\mu_1 \pm u_{22})]\cos\mu_1 \\ (\mp u_{11}\cos\mu_1 \pm u_{31}\sin\mu_1)m_{p1} - B_y \pm u_{12}\cos\mu_1 \pm u_{32}\sin\mu_1 \\ \mp[u_{21}m_{p1} - (B_z \sin\mu_1 - B_x \cos\mu_1 \pm u_{22})]\sin\mu_1 \end{bmatrix} \tag{34}$$

Vector $(\vec{\omega}^{(p)} x \vec{V}^{(1)})$ is represented by:

$$(\vec{\omega}^{(p)} x \vec{V}^{(1)}) = [\omega^{(1)}]^2 \begin{bmatrix} -(B_z \sin\mu_1 - B_x \cos\mu_1)m_{p1}\sin\delta_1 \\ -B_y m_{p1}\sin\gamma_1 \\ -(B_z \sin\mu_1 - B_x \cos\mu_1)m_{p1}\cos\delta_1 \end{bmatrix} \tag{35}$$

Subtracting **Eq.(35)** from **Eq.(34):**

$$\vec{\omega}^{(1)} x \vec{V}^{(p)} - \vec{\omega}^{(p)} x \vec{V}^{(1)} = [\omega^{(1)}]^2 \begin{bmatrix} h_{11}m_{p1} + h_{12} \\ h_{21}m_{p1} + h_{22} \\ h_{31}m_{p1} + h_{32} \end{bmatrix} \tag{36}$$

Where:

$$h_{11} = \pm u_{21}\cos\mu_1 - (B_x \cos\mu_1 - B_z \sin\mu_1)\sin\delta_1 \tag{37}$$

$$h_{12} = (B_z \sin\mu_1 - B_x \cos\mu_1 \pm u_{22})\cos\mu_1 \tag{38}$$

$$h_{21} = B_y \sin\gamma_1 \mp u_{11}\cos\mu_1 \pm u_{31}\sin\mu_1 \tag{39}$$

$$h_{22} = -(B_y \pm u_{12}\cos\mu_1 \pm u_{32}\sin\mu_1) \tag{40}$$

$$h_{31} = \mp u_{21} \sin\mu_1 - (B_x \cos\mu_1 - B_z \sin\mu_1)\cos\delta_1 \tag{41}$$

$$h_{32} = -(B_z \sin\mu_1 - B_x \cos\mu_1 \pm u_{22})\sin\mu_1 \tag{42}$$

Therefore, it may be obtained $\vec{n}.(\vec{\omega}^{(1)} x \vec{V}^{(p)} - \vec{\omega}^{(p)} x \vec{V}^{(1)})$ as follows:

$$\vec{n}.(\vec{\omega}^{(1)} x \vec{V}^{(p)} - \vec{\omega}^{(p)} x \vec{V}^{(1)}) = (f_1 m_{p1} + f_2)[\omega^{(1)}]^2 \tag{43}$$

Where:

$$f_1 = n_x h_{11} + n_y h_{21} + n_z h_{31} \tag{44}$$

$$f_2 = n_x h_{12} + n_y h_{22} + n_z h_{32} \tag{45}$$

Step 9: Representation of m_{p1}

Using **Eqs. (12)** and **(25)**, the equation of a_{13} may be represented by:

$$a_{13} = -(k_{PI} t_2 + C_{12})\omega^{(1)} \tag{46}$$

Using Eqs. (11) and (24), a23 may be described by:

$$a_{23} = -(k_{PII} t_4 + C_{22})\omega^{(1)} \tag{47}$$

Using **Eqs. (24), (25),** and **(43)**, the equation for a33 may be represented by:

$$a_{33} = [(2k_{PI} t_1 t_2 - \upsilon_2 - f_1)m_{p1} + (k_{PI} t_2^2 + k_{PII} t_4^2 - \upsilon_3 - f_2)][\omega^{(1)}]^2 \tag{48}$$

with:

$$a_{12} a_{33} - a_{13} a_{23} = 0 \tag{49}$$

So, **Eqs. (46), (47), (48),** and **(49)** yield:

$$m_{p1} = -\frac{a_{12}(k_{PI} t_2^2 + k_{PII} t_4^2 - \upsilon_3 - f_2) - (k_{PI} t_2 + C_{12})(k_{PII} t_4 + C_{22})}{a_{12}(2k_{PI} t_1 t_2 - \upsilon_2 - f_1)} \tag{50}$$

Determination of Parameters E_m and L_m

Parameters E_m and L_m of the pinion machine-tool settings have been shown in **Figs.(3)** and **(4).** Since the pinion cutting ratio m_{p1} has been determined, it is easy to find these two parameters **(Litvin et.al 2001).**

By using **Eq.(27)** to determine vector $\vec{V}^{(1p)}$ and applying **Eq.(16):**

$$E_m = \frac{\mp B_y \cos\mu_1 - V_x^{(1p)}}{m_{p1} \sin\delta_1} \mp B_y \tag{51}$$

Also:

$$L_m = \frac{B_y \cos\mu_1 - B_z \sin\mu_1 \mp V_y^{(1p)}}{m_{p1}} + B_x \sin\delta_1 + B_z \cos\delta_1 \tag{52}$$

Determination of Pinion Radial Setting and Cradle Angle

The determination of the pinion radial setting and the cradle angle is based on the consideration that the position vectors of the pinion tooth surface and head-cutter must coincide at the mean contact point.

For a straight-edged cutter(Lelkes and Marialigeti 2002)

$$s_p \sin q_p = \pm B_{fy} + E_m \pm u_p \sin\varphi_p \sin\tau_p \tag{53}$$

$$s_p \cos q_p = B_{fx} \sin\delta_1 + B_{fz} \cos\delta_1 - L_m - u_p \sin\varphi_p \cos\tau_p \tag{54}$$

For a curved-edged cutter (Theodore 1981)

$$s_p \sin q_p = \pm B_{fy} + E_m \pm \frac{\cos\tau_p \sin\tau_p}{k_{PI}} \tag{55}$$

$$s_p \cos q_p = B_{fx} \sin\delta_1 + B_{fz} \cos\delta_1 - L_m \pm \frac{\cos\tau_p \cos\tau_p}{k_{PI}} \tag{56}$$

Using $\sin^2 q_p + \cos^2 q_p = 1$, and eliminating q_p then solving for pinion radius s_p. Also eliminating s_p, it can be solved the pinion cradle angle q_p.

NUMERICAL EXAMPLE

The synthesis above is used to determine the machine-tool settings for a pair of spiral bevel gear drive, and then TCA computer program is developed to simulate the meshing of this pair under alignment and misalignment conditions, **Fig.(8)** represent the flow chart of the TCA computer program which has been developed in this work.

The major blank data is represented in **Table (1)**, these data imported from standard Gleason works **(Theodore 1981).**

The straight blade is used to cut the gears and curved blade is used to cut the pinion, and **Table (2)** shows the input design data. **Tables (3)** and **(4)** show the output for the pinion machine-tool settings.

Two conditions of misalignment are considered when the TCA is applied to simulate the meshing. They are the shift of pinion along its axis, which is denoted by ΔA, and the error of pinion shaft offset, which is denoted by ΔV.

The output of TCA program is shown in **Figs. (9 to 18)** and from these figures, its clear that there is a reduction in the transmission error for the case of using curved edged blade, where the value of the transmission error for the case of straight-edged blade with ΔA = − 0.05 mm, [**Fig.(10)**] is about (8 sec.) but for the case of using curved-edged blade with ΔA = − 0.05 mm, [Fig. (16)] is about (7 sec.), i.e. reduction ratio about (12% →15%). Also, the same tendency can be seen when

$\Delta A = +0.05$ mm, [**Fig. (10)** and **Fig. (15)**] and when $\Delta V = -0.05$ mm, [**Fig. (13)** and **Fig. (18)**], or $\Delta V = +0.05$ mm, [**Fig. (12)** and **Fig. (17)**].

CONCLUSION

The main conclusions obtained from present work can be summarized as follows:

1- A developed approach to simulate the optimal meshing and contact of spiral bevel gear drives has successfully applied by controlling the machine-tool settings.
2- A new approach of TCA has been proposed.
3- A computer program to evaluate the pinion machine-tool settings and function of transmission errors has been developed.
4- The results of this computer program show the effect procedure which followed and leads us for controlling the bearing contact by reducing the errors of misalignment.

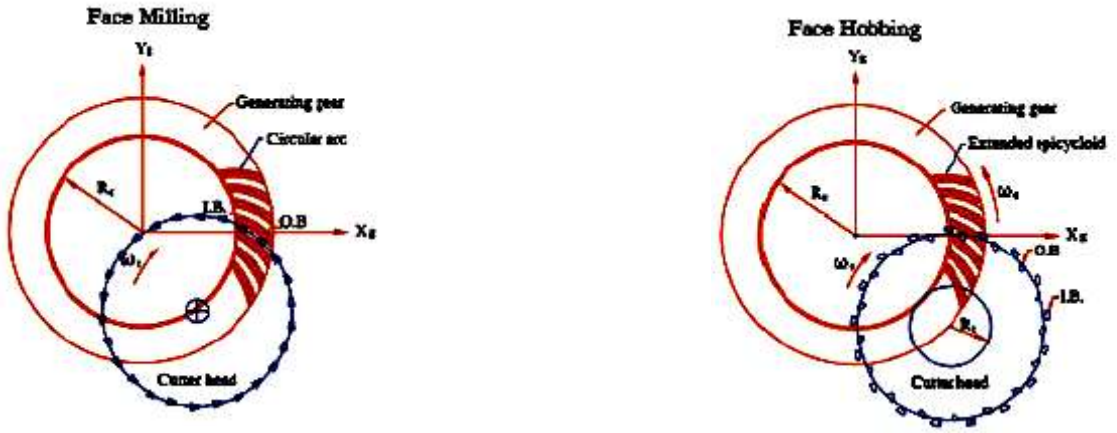


Fig.(1): Face Milling And Face Hobbing Generation Processes

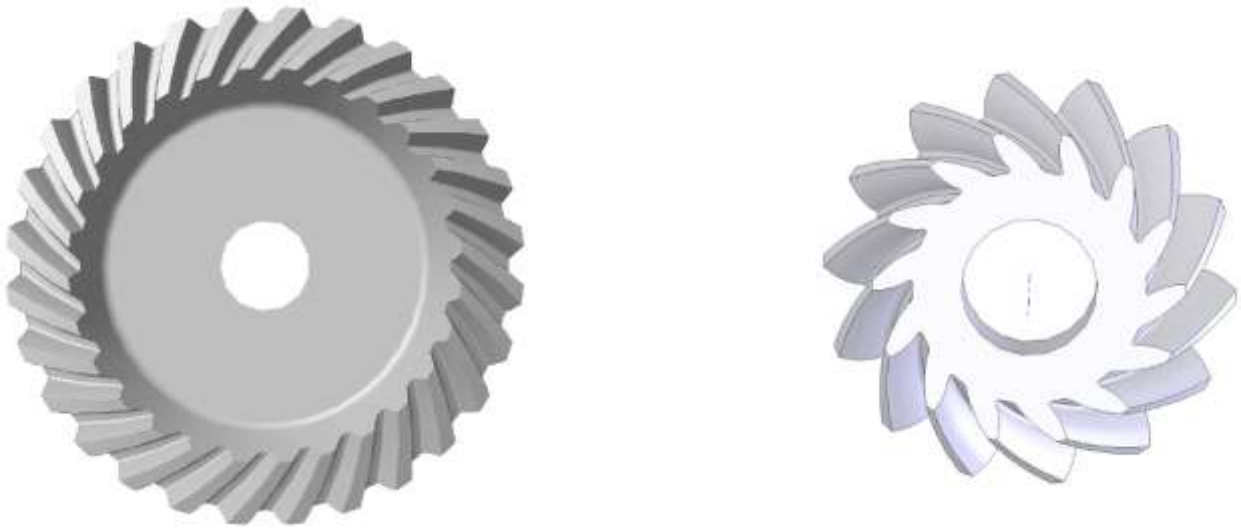


Fig.(2): 3D Geometric Model Of Gear And Pinion Created By Solid Works Program

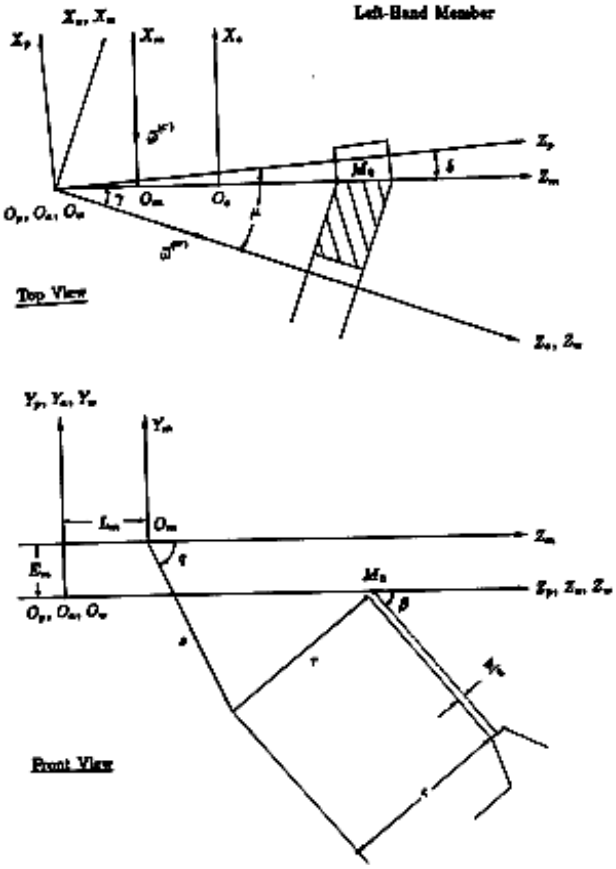


Fig. (3): Top And Front Views Of Left-Hand Gear Generator

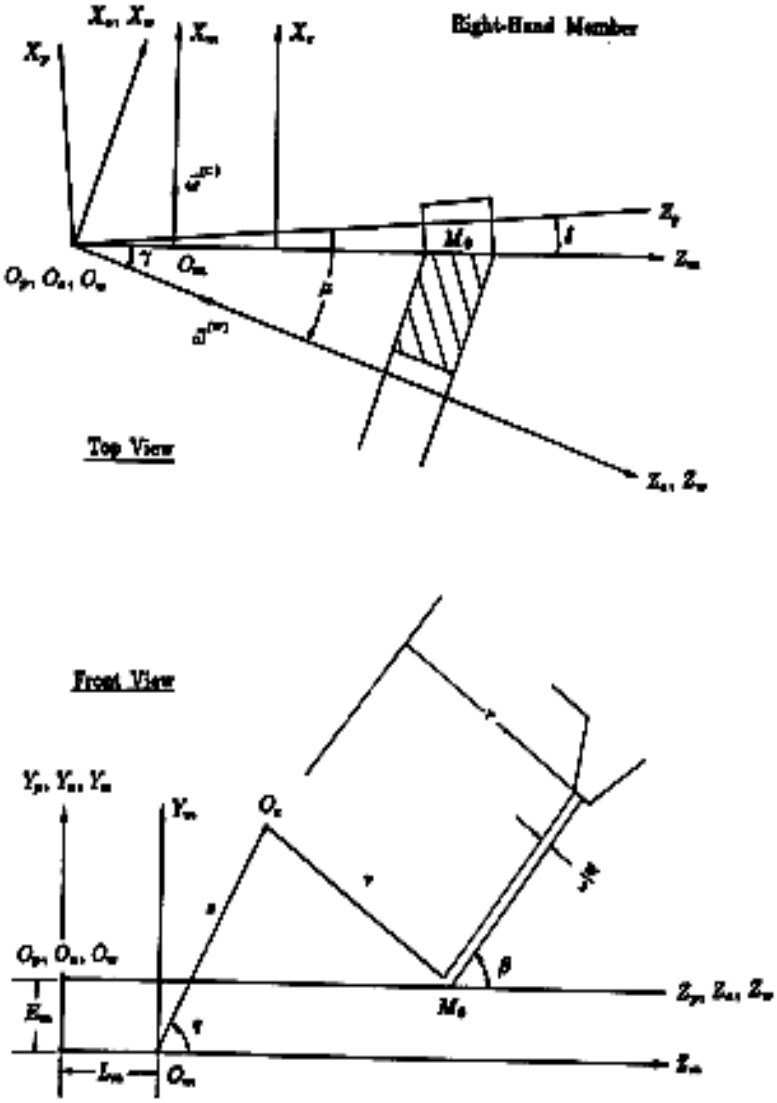


Fig. (4): Top And Front Views Of Right-Hand Gear Generator.

Left-Hand

Right-Hand

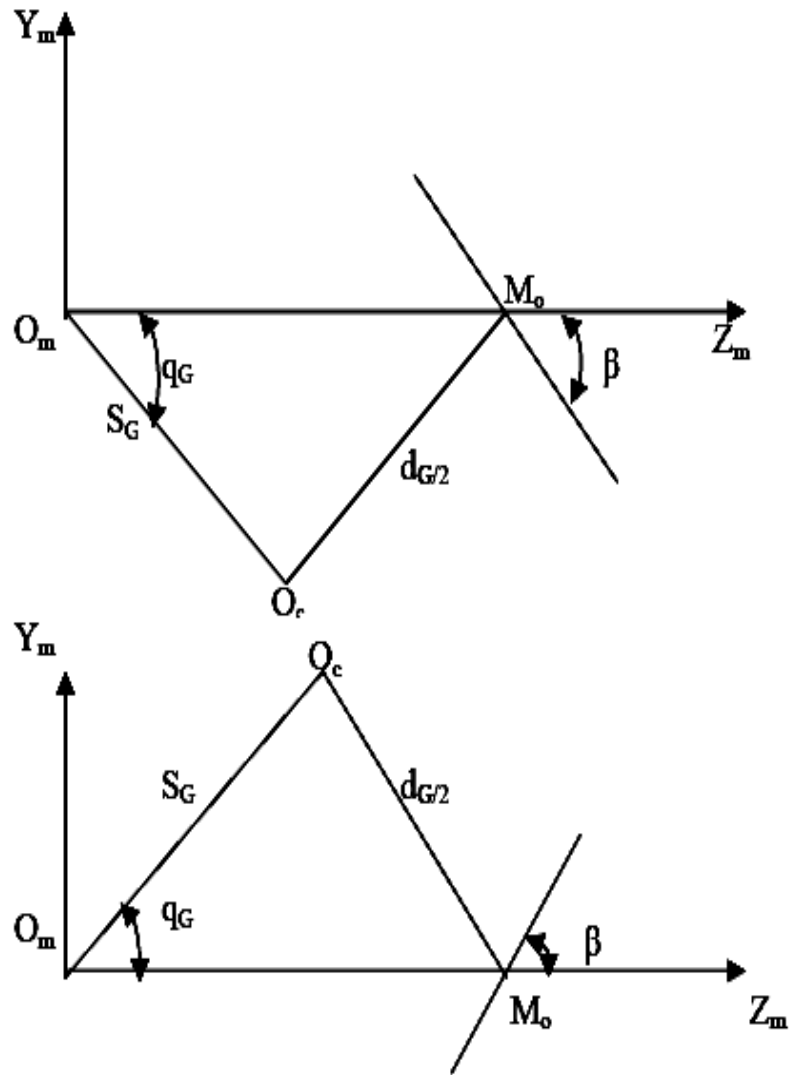


Fig.(5): The Front View Of The Installation Of The Head Cutter

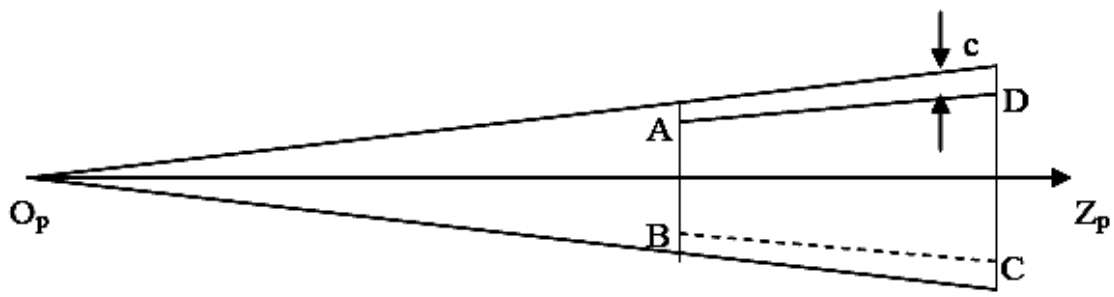


Fig. (6): Gear Tooth Surface

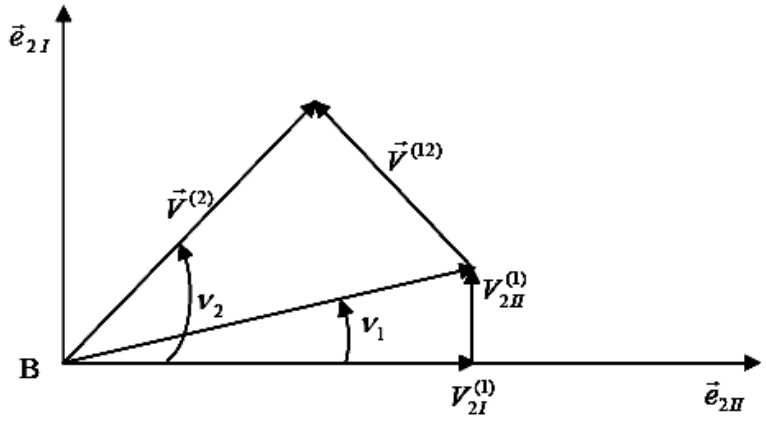


Fig.(7): Common Plane At The Mean Contact Point

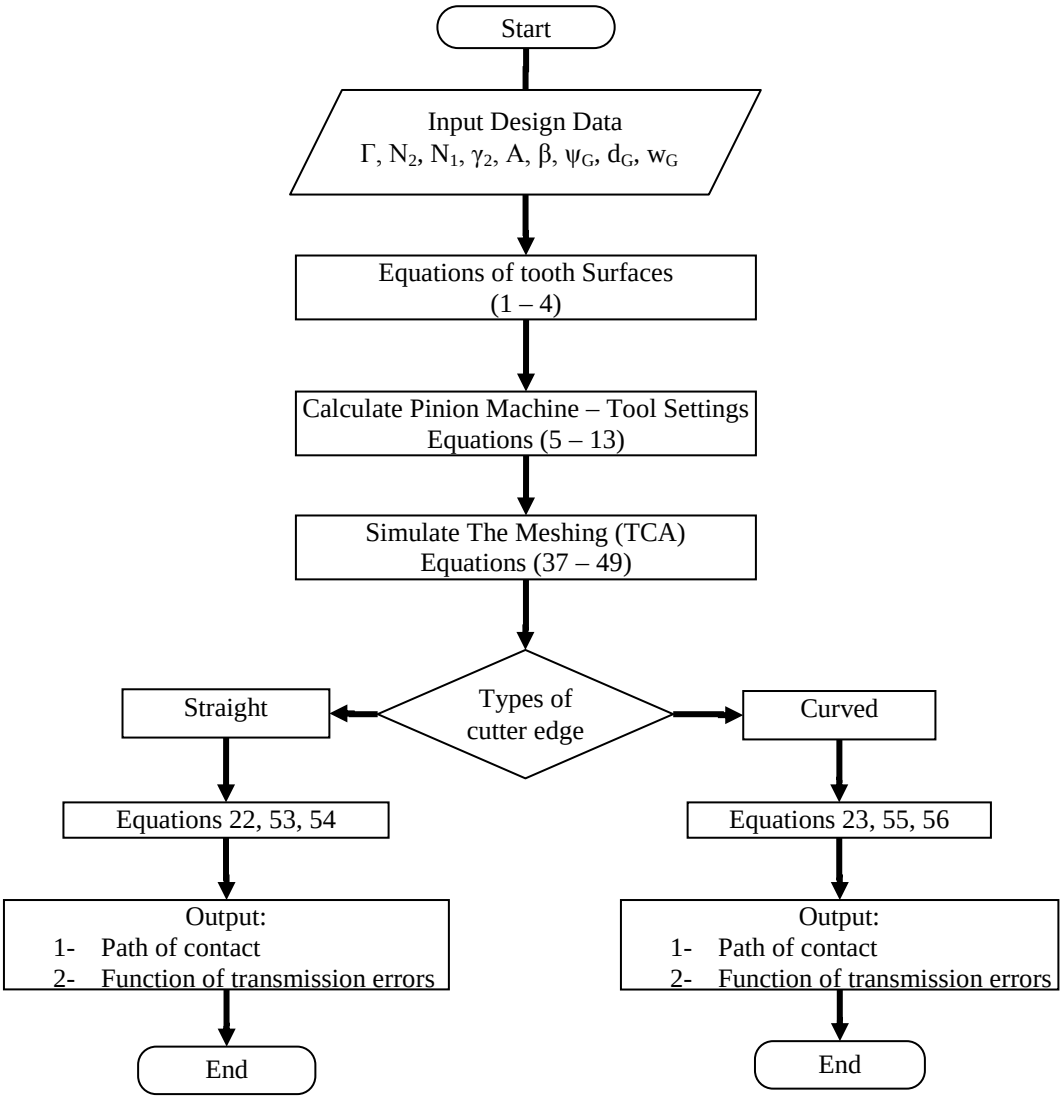


Fig.(8): Flow Chart Of Tooth Contact Analysis (TCA) Computer Program

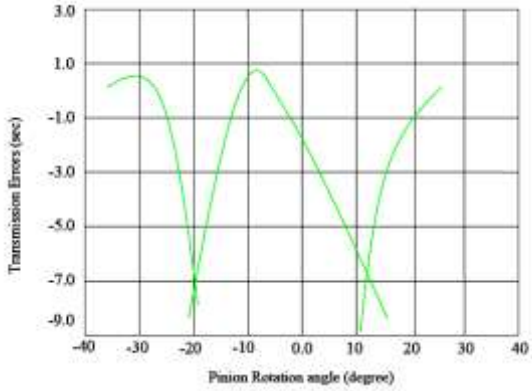


Fig.(9): TCA Output, Straight- Edged Blade, Alignment

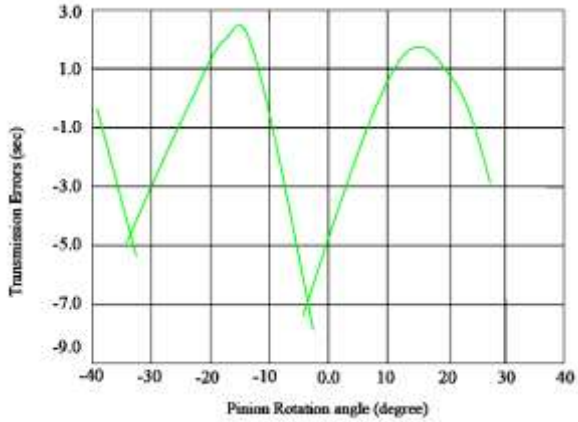


Fig.(10): TCA Output, Straight- Edged Blade, ΔA= +0.05 Mm

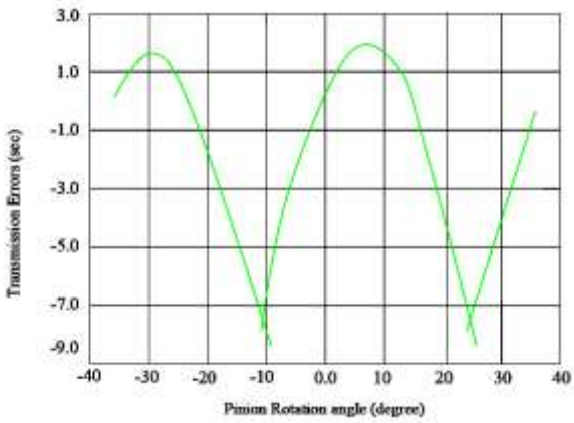


Fig.(11): TCA Output, Straight- Edged Blade, ΔA= -0.05 Mm

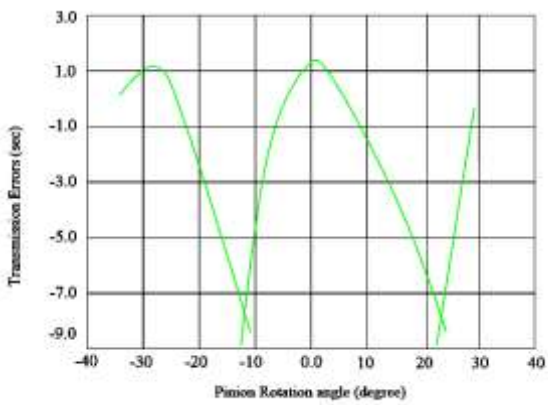


Fig.(12): TCA Output, Straight- Edged Blade, ΔV= +0.05 Mm

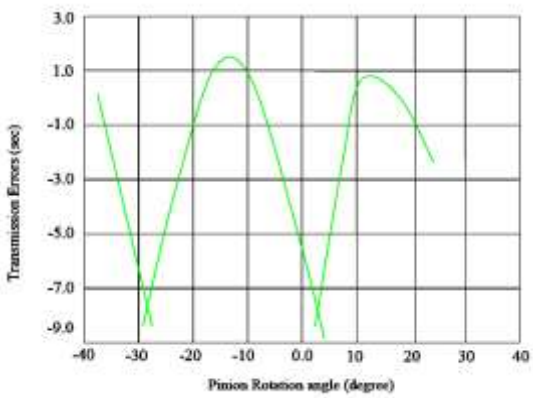


Fig.(13): TCA Output, Straight- Edged Blade, ΔV= -0.05 Mm

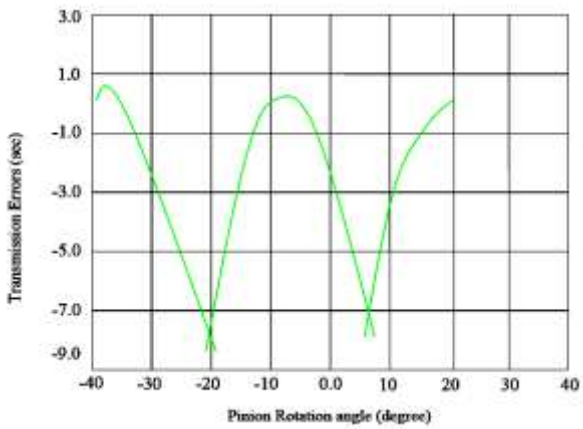


Fig.(14): TCA Output, Curved- Edged Blade, Alignment.

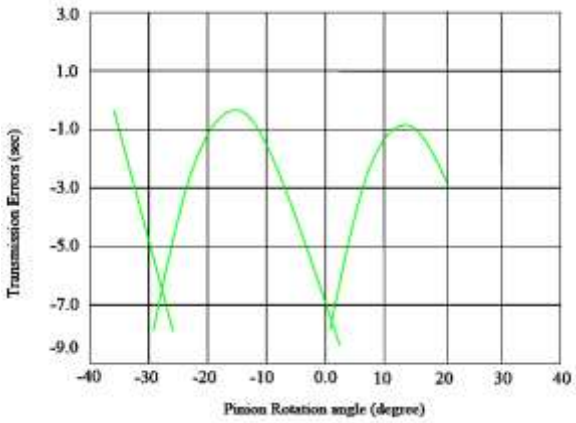


Fig.(15): TCA Output, Curved- Edged Blade, ΔA= +0.05

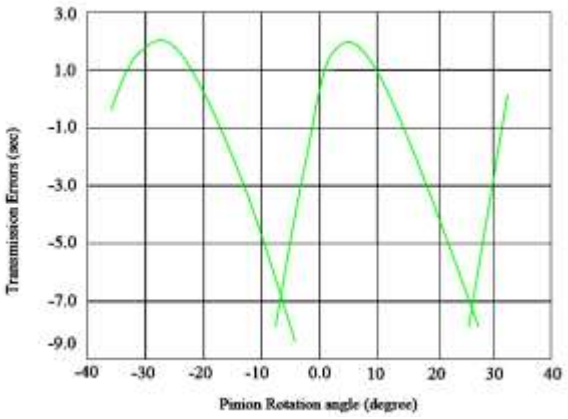


Fig.(16): TCA Output, Curved- Edged Blade, ΔA= -0.05

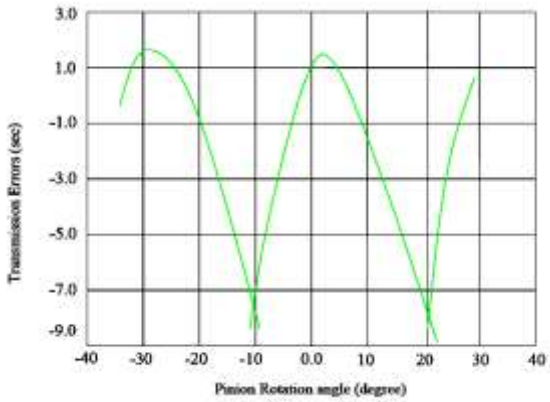


Fig.(17): TCA Output, Curved- Edged Blade, ΔV= +0.05

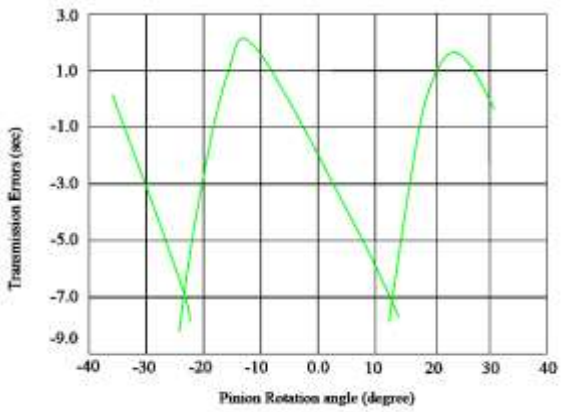


Fig.(18): TCA Output, Curved-Edged Blade, ΔV= -0.05

Table (1): Blank Data

	Pinion	Gear
Number of Teeth	10	41
Diametric Pitch	141.199 mm	
Shaft Angle	90°	
Mean Cone Distance	81.940 mm	
Outer Cone Distance	96.418 mm	
Whole Depth	8.509 mm	
Working Depth	7.671 mm	
Clearance	0.838 mm	
Face Width	28.931 mm	
Root Cone Angle	12°	72°
Mean Spiral Angle	35°	
Hand of Spiral	R.H	L.H

Table (2): Input Data

	Gear Convex Side	Gear Concave Side
Gear Blade Angle	20°	
Gear Cutter Average Diameter	152.399 mm	
Gear Cutter Point Width	2.032 mm	
First Derivative of Gear Ratio	- 0.0037	0.0055
Semi major Axis of Contact Ellipse	4.343 mm	4.343 mm
Contact Path Direction Angle	90°	75°
Radius of Blade	1016 mm	1270 mm

Table (3): Pinion Mation Settings With Straight Blade

	Pinion Concave Side	Pinion Convex Side
Blade Angle	16.5561^o	22.9907^o
Tip Radius of Cutter	75.303 mm	77.987 mm
Radial	76.030 mm	68.525 mm
Cradle Angle	63.1869^o	54.1910^o
Ratio of Roll	0.229	0.25348
Machining Offset	4.421 mm	- 6.213 mm
Machine Center to Back + Sliding Base	0.539 mm	1.324 mm

Table (4): Pinion Mation Settings With Curved Blade

	Pinion Concave Side	Pinion Convex Side
Blade Angle	16.5561^o	22.9907^o
Blade Center	(293.548, 0, -896.849) mm	(499.998, 0, 1244.752) mm
Tip Radius of Cutter	75.811 mm	77.314 mm
Radial	75.077 mm	69.662 mm
Cradle Angle	63.0025^o	54.09^o
Ratio of Roll	0.23157	0.24915
Machining Offset	3.059 mm	-4.782 mm
Machine Center to Back + Sliding Base	0.429 mm	0.916 mm

REFERENCES

Faydor L. Litvin and Alfonso Fuentes (Computerized design, Generation, Simulation of meshing and contact, and stress analysis of formate cut spiral bevel gear drives), NASA CR-525, 2003

Faydor L. Litvin and Alfonso Fuentes (Gear geometry and applied theory), Second Edition, Cambridge University Press 2004

Gosselin C, Cloutier L and Brousseau J. (Tooth contact analysis of high conformity spiral bevel gear), NASA CR-341, 1991

Joseph L. Arvin and Thomas C. (Spiral bevel gear development, eliminating trial and error with computer technology), Gear Technology, The Journal of gear manufacturing, February 2003

Lelkes M. and Marialigeti J. (Cutting definition for kinematic optimization of spiral bevel gears), Hungary 2002.

Litvin F.L and Lee H.(Generation and tooth contact analysis of spiral bevel gears with predesigned parabolic functions of transmission errors), NASA TR-C-014, 1989

Litvin F.L, Q. Fan, A. Fuentes and R.F Handschuh (Computerized design, generation, simulation of meshing and contact of face-milled formate-cut spiral bevel gears), NASA TR-1224, 2001

Qi Fan (Kinematical simulation of face Hobbing indexing and tooth surface generation of spiral bevel and Hypoid gears), AGMA Technichal Paper 91 FTM, 2006

Robert F. Handschuh (Recent advances in the analysis of spiral bevel gears), NASA ARL-TR-1316, 1997

Theodore J. Krenzer (TCA of spiral bevel gears under load), Gleason work, 1981

NOMENCLATUTE

Index

C	tool surface (C = G, P)
W	work surface (W = 1, 2)
G	gear tool surface
P	pinion tool surface
1	pinion surface
2	gear surface
I	first principal
II	second principal

Matrices and Vectors

[A]	matrix represents the relation between the principal curvatures and directions for mating surfaces
[B]	matrix represents homogenous coordinates of point B
$[L_{ab}]$	matrix describes the transformation of vector from the S_b coordinate system to S_a coordinate system
$[M_{ab}]$	matrix describes the transformation of coordinates from the S_b coordinate system to S_a coordinate system
[N]	matrix represents components of normal vector $\vec{N}$
[n]	matrix represents components of unit normal vector $\vec{n}$
[ω]	matrix represents components of angular velocity vector $\vec{\omega}$
$\vec{B}$	Position vector of point B on a surface
$\vec{B}_u$	$\partial\vec{B}/\partial u$
$\vec{B}_v$	$\partial\vec{B}/\partial v$
$\vec{e}_I, \vec{e}_{II}$	Unit vectors along the principal directions of the surface at the contact point
$\vec{i}, \vec{j}, \vec{k}$	Base vectors along axes X, Y, and Z, respectively
$\vec{N}$	Normal vector of point B on a surface

$\vec{n}$	Unit normal vector of point B on a surface
$\vec{n}_u$	$\partial\vec{n}/\partial u$
$\vec{n}_v$	$\partial\vec{n}/\partial v$
$\vec{V}^{(CW)}$	Slide velocity of surfaces ∑C and ∑W
$\vec{V}$	Transfer velocity vector
$\vec{V}^{(1)},\vec{V}^{(2)}$	Velocity vectors of contact point in its motion over the pinion and gear surfaces, respectively
$\vec{\omega}$	Angular velocity vector
$\vec{\omega}^{(TQ)}$	Relative angular velocity vector of surface T with respect to surface Q
$\vec{\tau}$	Tangent vector

Latin Symbols

A	mean pitch cone distance, mm
A_0,A_1,A_2	Coefficient of a quadratic equation
B	Point on a surface
E_m	Machining offset, mm
L_m	Vector sum of machine center to back and sliding base
m_{G2}	Gear cutting ratio
$V_{2I}^{(1)},V_{2II}^{(1)}$	The projections of vector $\vec{V}^{(1)}$ on vectors $\vec{e}_{2I}$ and $\vec{e}_{2II}$, respectively, mm/sec
X_{MCB}	Machine center to back
X_{SB}	Sliding base, mm
W	Point width, mm
a	Constant
a_{ij}	Element of matrix [A]
b	Semi-minor axis of the contact ellipse, mm
c	Clearance, mm
d_g	Average diameter of gear cutter, mm
k_n	Normal curvature, mm
q	Cradle angle, deg
r	Tip radius of the cutter, mm
s	Radial setting, mm
t	Semi-major axis of the contact ellipse, mm

Greek Symbols

∑	Surface
Γ	Shaft angle, deg
α	Orientation angle of ellipse, deg
β	Mean spiral angle, deg
δ	Dedendum angle, deg
ϵ	Specified tolerance value
γ	Root angle, deg
μ	Pitch angle, deg
ϖ^{G2}	Angular velocity in relative motion, rad/sec
υ_1,υ_2	Angles formed between vectors $\vec{V}^{(1)}$ and $\vec{e}_{2I}$, and $\vec{V}^{(2)}$ and $\vec{e}_{2I}$, respectively, deg.

Modified Stability Functions with Shear Effects for Non-Prismatic Members in Steel Plane Frames

Dr. Omar Al-Farouk Al-Damluji,
Assistant Professor,
Department of Civil Engineering,
University of Baghdad, Iraq.

Dr. Sabeeh Zaki AL-Sarraf,
Professor,
Department of Building and Construction,
University of Technology, Iraq.

Wisam Victor Yossif,
Formally Postgraduate Student,
Department of Civil Engineering,
University of Baghdad, Iraq.

ABSTRACT

The mathematical model of the tapered struts subjected to axial load is solved to obtain the modified stability functions due to shear effect as well as bending effects. The stability functions are derived for a wide range of non-prismatic struts then compared in graphical curves with stability functions excluding shear effects.

The stability functions for non-prismatic members under compressive and tensile axial loads are developed for the purpose of expressing both effects of bending and shear in a beam-column stiffness at any value of axial force under the buckling limit.

الخلاصة

لقد تم حل نموذج رياضي لعتبات لا موشورية معرضة لثقل محوري للحصول على دوال استقرارية معدلة ناتجة عن تاثيرات القص اضافة الى تاثيرات الانحناء المعهودة. لقد تم اشتقاق دوال الاستقرارية لمدى واسع من العتبات اللاموشورية ومن ثم تمت المقارنة بمنحنيات مرتسمية مع دوال الاستقرارية من دون تاثيرات القص.

لقد تم تطوير دوال الاستقرارية للاعضاء اللاموشورية تحت اثقال انضغاط و شد محورية في سبيل التعبير عن تاثيري الانحناء و القص في جساءة العتبة–العمود تحت اية قيمة من قوة محورية اقل من حد الانبعاج.

KEYWORDS

Modified Stability Functions – Shear Effects – Non-Prismatic Members – Beam-Column Stiffness

INTRODUCTION

Shear effects on the elastic stability analysis are usually not considered in the analysis and design of frames made up of structural members of solid sections; this is because the distortion caused by shear is relatively small except for short members. But in different member's length and cross sectional shapes, the contribution of shear deformation to the total deflection may be appreciable. Few research works had covered shear effects in the non-

prismatic members loaded axially. Most of the researchers used approximate methods for including this effect such as a method proposed by Al-Quraishi [1] which used approximate stability functions for non-prismatic members by developing the approximate formula similarly to that derived by S. Al-Sarraf[3]. Other researchers such as AL-Fadul[2] obtained the exact stability functions for special shape of non-prismatic members and used the finite difference method to obtain the approximate stability functions for other shapes.

In the present study, a new method is adopted based on the exact stability functions including bending effects. The new expression of stability functions is derived herein by adding the effect of shear to the slope of deflection curve and to the member curvature.

Mathematical Model

When the beam-column member is loaded with a constant axial force Q, the initially straight longitudinal axis is deformed into a curve called the deflection curve which is produced by the combined effects of bending and shear deformations.

The ratio of change in slope caused by shear to that caused by moment is defined as the *shear flexibility parameter* as given by Equation (1) for prismatic members [3, 4, 5], Equation (2) for batten lacing prismatic members [1, 6] and Equation (3) for non-prismatic members [1]. The modified stability functions for beam-columns having any solid cross-sectional shapes, built-up structural members are developed in terms of the shear flexibility parameter.

$$\mu = \frac{Q_E}{GA_V} \tag{1}$$

$$\mu = Q_E \cdot \left(\frac{l_v . l_h}{12EI_b} + \frac{l_v^{\,2}}{24EI_b} + \frac{n.l_v}{l_h . GA_b} \right) \tag{2}$$

$$\mu_2 = \frac{Q_{E2}}{GA_{V2}} \tag{3}$$

where

$Q_E = \frac{\pi^2 EI}{L^2}$: Euler load for prismatic members,or

$Q_E = \frac{\pi^2 EI_c}{L^2}$: Euler load for batten prismatic members,

$Q_{E2} = \frac{\pi^2 EI_2}{L^2}$: Euler load for tapered members at end 2,

EI, GA_v : flexural and shear rigidities of prismatic members ,

EI_b, GA_b : flexural and shear rigidities of batten prismatic members ,

EI_2, GA_{v2} : flexural and shear rigidities of tapered members at end 2,

I_c : moment of inertia for the vertical plate,

A_v : effective shear area for prismatic members,

A_b : cross sectional area for batten,

A_{v2} : effective shear area for tapered members at end 2,

n : numerical factor equal to 1.2 in the case of a rectangular cross section[7],

l_h : length of the batten center to center of the vertical plate, and

l_v : center to center vertical distance between two batten lacing.

PROPOSED MODIFIED STABILITY FUNCTION FOR NON-PRISMATIC MEMBERS:

A member may have a linear taper in either one or the other direction as shown in Figure (1), M_1 and M_2 are the applied end moments, a is the distance of end 2 from the origin O and b is the distance of end 1 from the origin O of zero depth. Therefore, the depth d(x) may be expressed by Equation (4):

$$d(x) = d_2(x/a) \tag{4}$$

where

d_1, d_2 : depths of member at ends 1and 2, respectively,

d(x) : depth of member at variable (x) from origin (O).

The moment of inertia of the member about the axis of bending is expressed in the form:

$$I(x) = I_2(x/a)^m \tag{5}$$

where I(x) is the moment of inertia at distance x from the origin O, m is the shape factor that depends on the cross-sectional shape as given in Table (1).

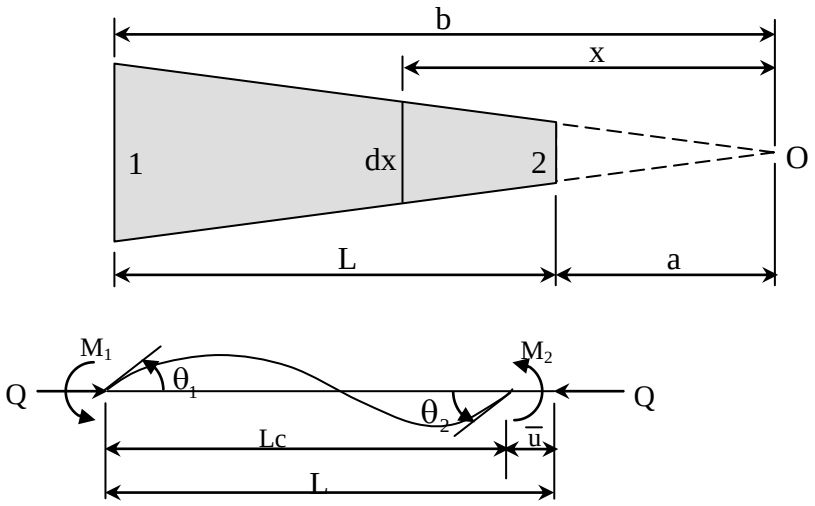


Fig (1): Non-prismatic beam-column element

The effect of shear force of the non-prismatic beam-column is derived starting from the slope of the deflection curve [10]:

$$\frac{dy}{dx} = \frac{dy_b}{dx} + \frac{V(x)}{Av(x).G} \tag{6}$$

$$\frac{d^2y}{dx^2} = -\frac{M(x)}{EI(x)} + \frac{d}{dx}\left(\frac{V(x)}{Av(x).G}\right) \tag{7}$$

The bending moment and the shear force at distance x from the origin are: -

$$M(x) = \left(\frac{a-x}{L}\right)M_1 + \left(\frac{b-x}{L}\right)M_2 \mp Q(y_b + y_s) \tag{8}$$

$$V(x) = -\left(\frac{M_1 + M_2}{L}\right) + \left(Q\frac{dy_b}{dx} + Q\frac{dy_s}{dx}\right) \tag{9}$$

By substituting Equation (9) into the slope due to shear yields: -

$$\frac{dy_s}{dx} = \frac{1}{Av(x).G}.\left[-\left(\frac{M_1 + M_2}{L}\right) + Q\frac{dy_s}{dx}\right] \tag{10}$$

The effective area under shear stress at distance (x) from the origin is: -

$$Av(x) = Av_2\left(\frac{x}{a}\right)^n \tag{11}$$

The shear flexibility parameter and non-dimensional axial force parameter can be defined by the equation below:

$$\frac{Q}{Av_2.G} = \mu_2.\rho_2 \tag{12}$$

The curvature equation due to shear is derived below: -

$$\frac{d^2y_s}{dx^2} - n.\left(\frac{M_1+M_2}{L}\right).\frac{a^n}{x}.\frac{1}{\left(1-\mu_2.\rho_2.\frac{a^n}{x^n}\right)^2}.\frac{\mu_2.\rho_2}{Q} = 0 \tag{13}$$

MODIFIED STABILITY FUNCTIONS FOR BEAM-COLUMN WITH SQUARE CROSS SECTIONS

The basic differential equation when the shape factor n = 2 is [10]:

$$\frac{d^2y_s}{dx^2} - 2.\left(\frac{M_1+M_2}{L}\right).\frac{a^2}{x}.\frac{1.}{\left(1-\mu_2.\rho_2.\frac{a^2}{x^2}\right)^2}.\frac{\mu_2.\rho_2}{Q} = 0 \tag{14}$$

The solution of the second order differential equation is: -

$$y_s = \left(\frac{M_1+M_2}{L}\right).\frac{\sqrt{\mu_2.\rho_2.a^2}}{Q}.a\tanh\left(\frac{x}{\sqrt{\mu_2.\rho_2.a^2}}\right) + C1.x + C2 \tag{15}$$

where C1 and C2 are constants of integration obtained by substituting the boundary conditions:

y = 0 at x = a and x = b yields:

$$C1 = \left(\frac{M_1+M_2}{L}\right).\frac{\sqrt{\mu_2.\rho_2.a^2}}{Q}.\frac{a\tanh\left(\frac{a}{\sqrt{\mu_2.\rho_2.a^2}}\right) - a\tanh\left(\frac{b}{\sqrt{\mu_2.\rho_2.a^2}}\right)}{L} \tag{16}$$

$$C2 = \pm\left(\frac{M_1+M_2}{L}\right).\frac{\sqrt{\mu_2.\rho_2.a^2}}{Q}.\frac{a\tanh\left(\frac{b}{\sqrt{\mu_2.\rho_2.a^2}}\right).a - a\tanh\left(\frac{a}{\sqrt{\mu_2.\rho_2.a^2}}\right).b}{L} \tag{17}$$

The slope of the deflection curve due to shear only is: -

$$\frac{dy_s}{dx} = -\left(\frac{M_1+M_2}{L}\right).\frac{a^n}{x^n}.\left(1\pm\mu_2.\rho_2.\frac{a^2}{x^2}\right)^{-1}.\frac{\mu_2.\rho_2}{Q} + C1 \tag{18}$$

The curvature of the deflected curve due to bending only is: -

$$\frac{d^2y_b}{dx^2} = -\frac{M(x)}{EI(x)} \tag{19}$$

$$I(x) = I_2\left(\frac{x}{a}\right)^m \tag{20}$$

By substituting Equation (8) and (20) into Equation (19), yields: -

$$EI_2\left(\frac{x}{a}\right)^4\frac{d^2y_b}{dx^2}+Qy_b=\frac{M_1}{L}(x-a)+\frac{M_2}{L}(x-b) \qquad (21)$$

where $\omega^2=\left[Qa^4/E\,I_2\right]^{0.5}$.

The solution of Equation (21) and its first derivative are:

$$y_b=\sqrt{x}\left[AJ_{1/2}\left(\frac{\omega}{x}\right)+BJ_{-1/2}\left(\frac{\omega}{x}\right)\right]+\frac{M_1}{QL}(x-a)+\frac{M_2}{QL}(x-b) \qquad (22)$$

$$\frac{dy_b}{dx}=A\left(\frac{\omega}{x^{1.5}}\right)J_{\frac{3}{2}}\left(\frac{\omega}{x}\right)-B\left(\frac{\omega}{x^{1.5}}\right)J_{-\frac{3}{2}}\left(\frac{\omega}{x}\right)+\frac{M_1+M_2}{QL} \qquad (23)$$

where the boundary conditions are:

at x = a, y = 0, and dy/dx = θ_{2b},

and at x = b, y = 0, and dy/dx = θ_{1b}.

The values of the constants A and B are obtained by substituting the boundary conditions into Equation (22).

$$A=\frac{M_1J_{-\frac{1}{2}}(\alpha)\sqrt{a}+M_2J_{-\frac{1}{2}}(\beta)\sqrt{b}}{\sqrt{a}\,\sqrt{b}\;Z\,Q}$$

$$B=-\frac{M_1J_{\frac{1}{2}}(\alpha)\sqrt{a}+M_2J_{\frac{1}{2}}(\beta)\sqrt{b}}{\sqrt{a}\,\sqrt{b}\;Z\,Q}$$

where

$$Z=J_{\frac{1}{2}}(\alpha)J_{-\frac{1}{2}}(\beta)-J_{-\frac{1}{2}}(\alpha)J_{\frac{1}{2}}(\beta)$$

$$\alpha=\frac{\omega}{a}\;,\;\beta=\frac{\omega}{b}\;,\;\rho_2=\frac{QL^2}{EI_2\pi^2}\;,\;\omega=\left(\frac{a^4\,Q}{E\,I_2}\right)^{0.5}$$

Therefore, the total slope of the deflected curve which is $\frac{dy}{dx}=\frac{dy_b}{dx}+\frac{dy_s}{dx}$ becomes:

$$\frac{dy}{dx}=A\left(\frac{\omega}{x^{1.5}}\right)J_{\frac{3}{2}}\left(\frac{\omega}{x}\right)-B\left(\frac{\omega}{x^{1.5}}\right)J_{-\frac{3}{2}}\left(\frac{\omega}{x}\right)+\frac{M_1+M_2}{QL}-\left(\frac{M_1+M_2}{L}\right).\frac{a^2}{x^2}.\frac{1}{\left(1-\mu_2.\rho_2.\frac{a^2}{x^2}\right)}.\frac{\mu_2.\rho_2}{Q}+C1 \qquad (24)$$

The boundary conditions of $\frac{dy}{dx}$ at x = a and at x = b are:

$$\frac{dy}{dx}=\frac{\theta_{1b}-\frac{\mu_2.\rho_2}{Q.u^2}\left(\frac{M_1+M_2}{L}\right)}{1-\frac{\mu_2.\rho_2}{u^2}} \qquad \text{at } x=b \qquad (25)$$

$$\frac{dy}{dx}=\frac{\theta_{2b}-\frac{\mu_2.\rho_2}{Q}\left(\frac{M_1+M_2}{L}\right)}{1-\mu_2.\rho_2} \qquad \text{at } x=a, \qquad (26)$$

By substituting the above boundary conditions in Equation (24), then:

$$\left(A\left(\frac{\omega}{b^{1.5}}\right)J_{\frac{3}{2}}\left(\frac{\omega}{b}\right)-B\left(\frac{\omega}{b^{1.5}}\right)J_{-\frac{3}{2}}\left(\frac{\omega}{b}\right)+\frac{M_1+M_2}{Q.L}-\left(\frac{M_1+M_2}{Q.L}\right).\frac{\mu_2.\rho_2}{u^2-\mu_2.\rho_2}+C1\right).\left(1-\frac{\mu_2.\rho_2}{u^2}\right)$$

$$+\frac{\mu_2.\rho_2}{Q.u^2}\left(\frac{M_1+M_2}{L}\right)=\theta_{1b} \qquad \text{at } x = b \qquad (27)$$

$$\left(A\left(\frac{\omega}{a^{1.5}}\right)J_{\frac{3}{2}}\left(\frac{\omega}{a}\right)-B\left(\frac{\omega}{a^{1.5}}\right)J_{-\frac{3}{2}}\left(\frac{\omega}{a}\right)+\frac{M_1+M_2}{Q.L}-\left(\frac{M_1+M_2}{Q.L}\right).\frac{\mu_2.\rho_2}{1-\mu_2.\rho_2}+C1\right).(1-\mu_2.\rho_2)$$

$$+\frac{\mu_2.\rho_2}{Q}\left(\frac{M_1+M_2}{L}\right)=\theta_{2b} \qquad \text{at } x = a \qquad (28)$$

Thus the stability functions are:

$$S_1=\frac{Z_1}{EI_2}\frac{K_a(D_1QL+B_1QL+1)+\mu_2.\rho_2}{Q} \qquad (29)$$

$$SC=\frac{Z_1}{EI_2}\frac{K_b u^2(C_1QL+A_2QL+1)+\mu_2.\rho_2}{Qu^2} \qquad (30)$$

$$S_2=\frac{Z_1}{EI_2}\frac{K_b u^2(A_1QL+A_2QL+1)+\mu_2.\rho_2}{Q.u^2} \qquad (31)$$

where the symbols A1, B1, C1, D1, A2, B2, Ka, Kb and Z are given in Appendix

The modified stability functions for beam-column with a rectangular cross section bent about the major axis and a rectangular cross section bent about the minor axis are listed in the Appendix:

Experimental Determination of the Elastic Critical Load:

It is based on the fact that the stiffness of a structure decreases with the increase of the axial force in the members. At the critical load, the stiffness of the structure is zero. Thus, it is possible to determine the elastic critical load of structures by expressing the stiffness as described below.

The natural frequency of oscillation of a frame decreases with the increase of the applied external load. When the vibration stiffness is plotted against the applied load, an almost linear relationship is obtained. An estimation of the elastic critical load is made by the extrapolation of the linear part of the graph. In frames with complicated sways, experimental determination of the elastic critical load is obtained by dynamic stiffness plots.

In general, for any system, the frequency of vibration may be calculated as in the following [8]: -

$$f=\frac{1}{2\pi}\sqrt{\frac{k}{\frac{\Sigma W}{g}}} \qquad (32)$$

where

f : frequency in cycle/sec,
k : stiffness of structure,
W : summation of external load on frame, and
g : acceleration due to gravity.

Equation (32) may be written as: -

$$k = \frac{(2\pi)^2}{g} \Sigma W f^2 \quad \text{for constant g,} \qquad (33)$$

$$k \propto \Sigma W.f^2$$

The stiffness of the primary buckling mode is a linear function of the applied external load. Thus, the relation between the total applied load ΣW and the parameter $(\Sigma W f^2)$ is approximately linear.

This method was widely used in finding out the elastic critical load experimentally. In this test which is done in present study, the structure is subjected to push at every applied load and the number of vibrations is calculated at this time. Then, the frequency, f, is equal to the number of vibrations over the unit time for that total load.

The cantilever beam-column models are manufactured from steel materials with dimensions of high accuracy. These models consist of four different non-prismatic cross-sectional types and one prismatic type in four different cases of batten lacing. The number of sway cycles is recorded for 6 seconds for different loads at the free end, then the relation between the loads and the beam-column dynamic stiffness is drawn to obtain the elastic critical load. The results of each type of non-prismatic member that is subjected to experimental load are reported in Table (1). The same thing is used for each case of batten lacing. The model types are shown in Appendix. Table (2) to (9) expose the average recorded data for three non-prismatic models of the described properties in Table (1). In other words, each type has dimensions defined in Table (1).

Table (1) Models dimensions

Test No.	Cross-section	m	u	Length, cm	Width, mm	Depth, mm
1	Square	4	2	30	6 – 12	6 – 12
2	Rectangular	3	2	40	3	1.5 – 3
3	Rectangular	1	2	40	15 – 30	2
4	Rectangular box 1 mm thickness	2.4	2	75	1.5	7 – 14
5	2 cm Batten lacing	-	1	40	25	2mm each part
6	5 cm Batten lacing	-	1	40	25	2mm each part
7	10 cm Batten lacing	-	1	40	25	2mm each part
8	20 cm Batten lacing	-	1	40	25	2mm each part

Table (2): Average results of 3 models having shape factor m=4

Axial Force, kN	Mass Load, kg	No. of cycles	Frequency rad/6 sec
0.0004905	0.050	49	51.31268
0.0009810	0.100	32	33.51032
0.0014715	0.150	23	24.08554
0.0019620	0.200	18	18.84956
0.0024525	0.250	14	14.66077
0.0029430	0.300	11	11.51917

Table (3): Average results of 3 models having shape factor m=3

Axial Force, kN	Mass Load, kg	No. of cycles	Frequency, rad/6 sec
0.000981	0.100	155	162.3156
0.001962	0.200	109	114.1445
0.002943	0.300	89	93.20059
0.003924	0.400	76	79.58702
0.004905	0.500	68	71.20944
0.009810	1.000	47	49.21829
0.049050	5.000	16	16.75516

Table (4): Average results of 3 models having shape factor m=1

Axial Force, kN	Mass Load, kg	No. of cycles	Frequency rad/6 sec
0.000981	0.100	84	87.9645
0.001962	0.200	59	61.7846
0.002943	0.300	47	49.2182
0.003924	0.400	41	42.9350
0.004905	0.500	36	37.6990
0.009810	1.000	24	25.1327
0.049050	5.000	3	3.1415

Table (5): Average results of 3 models having shape factor m=2.4

Axial Force, kN	Mass Load kg	No. of cycles	Frequency, rad/6 sec
0.000981	0.100	292	305.7817
0.001962	0.200	206	215.7227
0.002943	0.300	168	175.9292
0.003924	0.400	146	152.8908
0.004905	0.500	130	136.1357
0.009810	1.000	92	96.34217
0.049050	5.000	40	41.8879

Table (6): Average results of 3 prismatic models with 2 cm batten lacing

Axial Force, kN	Mass Load, kg	No. of cycles	Frequency rad/6 sec
0.000981	0.100	1007	1054.528
0.001962	0.200	712	745.604
0.002943	0.300	581	608.421

0.003924	0.400	503	526.740
0.004905	0.500	450	471.238
0.009810	1.000	318	333.008
0.098100	10.00	100	104.719

Table (7): Average results of 3 prismatic models with 5 cm batten lacing

Axial Force, kN	Mass Load, kg	No. of cycles	Frequency rad/6 sec
0.000981	0.100	603	631.460
0.001962	0.200	426	446.106
0.002943	0.300	348	364.424
0.003924	0.400	301	315.206
0.004905	0.500	269	281.696
0.009810	1.000	190	198.967
0.049050	5.000	85	89.011
0.098100	10.00	60	62.831

Table (8): Average results of 3 prismatic models with 10 cm Batten lacing

Axial Force, kN	Mass Load, kg	No. of cycles	Frequency, rad/6 sec
0.000981	0.100	356	372.8023
0.001962	0.200	252	263.8938
0.002943	0.300	205	214.6755
0.003924	0.400	178	186.4012
0.004905	0.500	159	166.5044
0.009810	1.000	112	117.2861
0.049050	5.000	50	52.35988
0.098100	10.00	35	36.65191

Table (9): Average results of 3 prismatic models with 20 cm batten lacing

Axial Force, kN	Mass Load, kg	No. of cycles	Frequency, rad/6 sec
0.000981	0.100	194	203.1563
0.001962	0.200	137	143.4661
0.002943	0.300	112	117.2861
0.003924	0.400	97	101.5782
0.004905	0.500	87	91.10619
0.009810	1.000	61	63.87905
0.049050	5.000	27	28.27433
0.098100	10.00	19	19.89675

The results are obtained experimentally and theoretically. They are then compared as in the following:

Experimentally: The results are analyzed graphically to obtain the buckling load at the vanished stiffness by extending the relation between the stiffness and mass linearly. Figures (2) to (9) show the buckling load for eight different models.

Theoretically: For a cantilever tapered column having load at the top free end, the stiffness matrix is[9]:

$$k = \begin{bmatrix} S_r & -(S_r + SC) \\ -(S_r + SC) & q \end{bmatrix}$$

where $q = S_1 + S_2 + 2SC - \pi^2 \rho_2$, $\rho_2 = \dfrac{QL^2}{EI_2 \pi^2}$

The results shown in Tables (10) and (11) present the obtained results by considering the eight different models experimentally and theoretically.

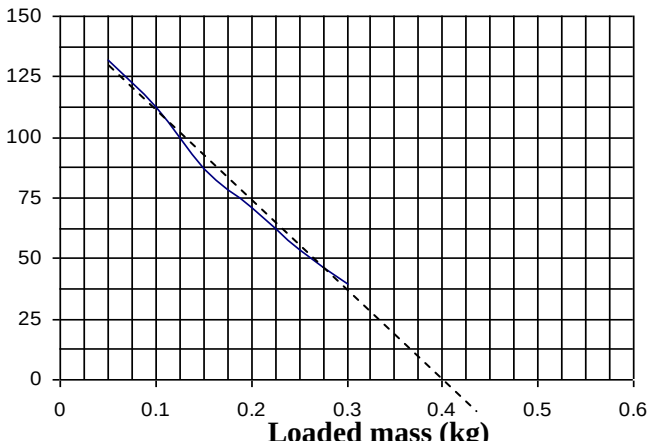


Fig (2): Stiffness-mass relation of model 1

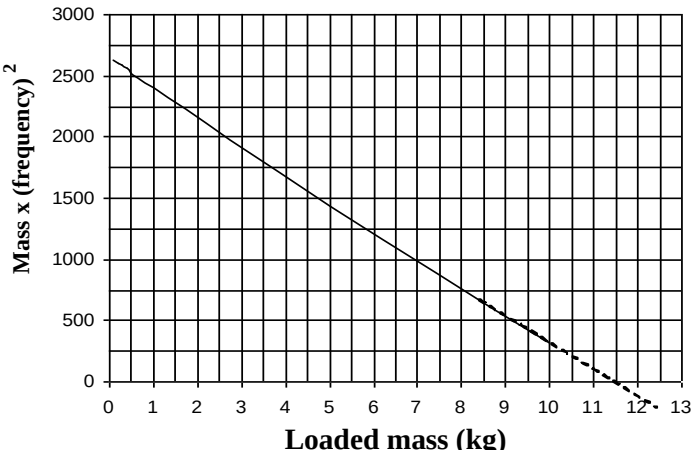


Fig (3): Stiffness-mass relation of model 2

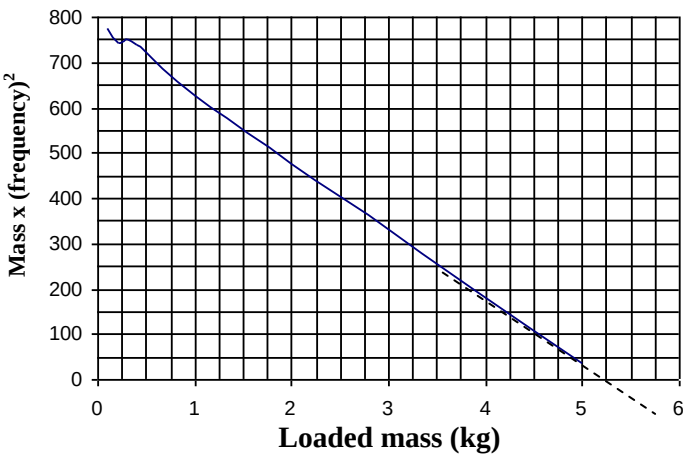


Fig (4): Stiffness-mass relation of model 3

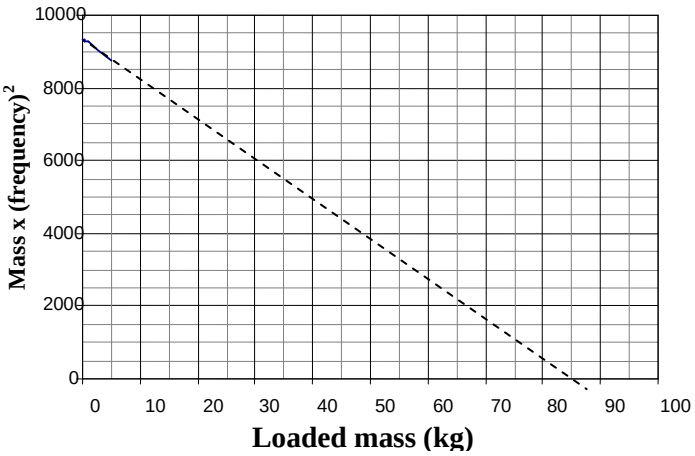


Fig (5): Stiffness-mass relation of model 4

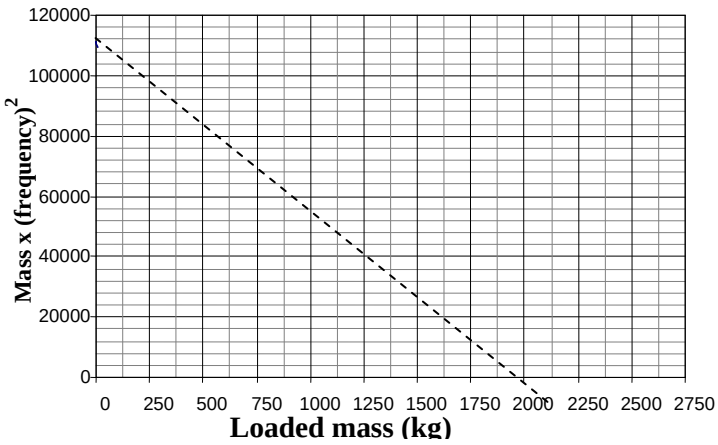


Fig (6): Stiffness-mass relation of model 5

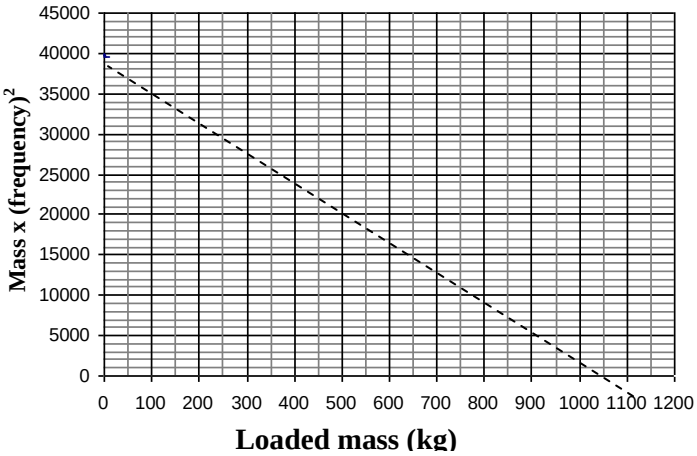


Fig (7): Stiffness-mass relation of model 6

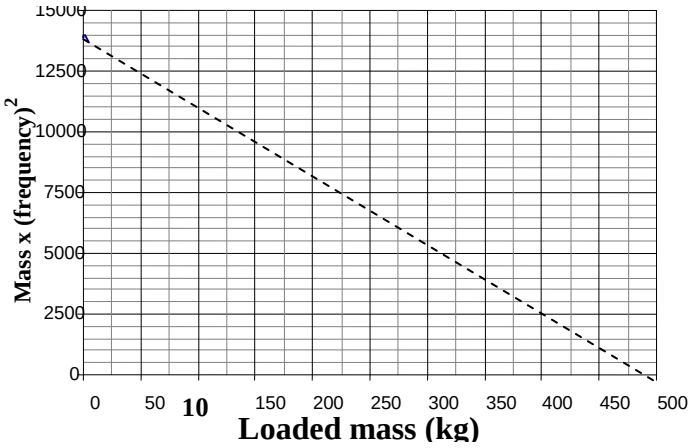


Fig (8): Stiffness-mass relation of model 7

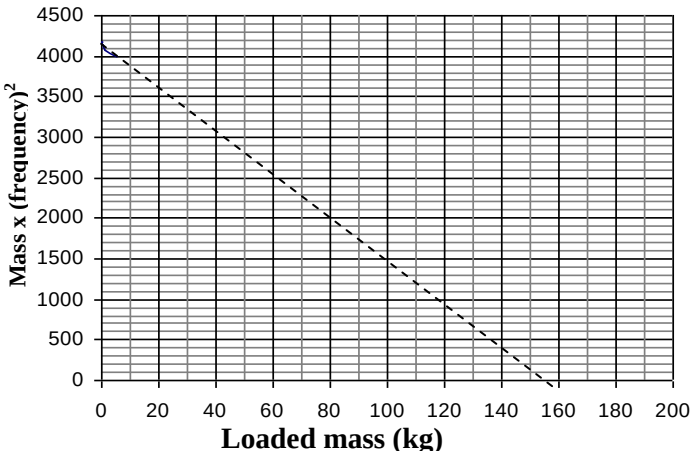


Fig (9): Stiffness-mass relation of model 8

(10) Experimental and theoretical results of model 1, 2, 3 and 4

Test No.	Experimental results		Theoretical results			
	$Q_{cr,}$ kN	$Q_{cr,}$ kg	ρ_2	S_1	S_2	SC
1	0.003946	0.405	1.6659	27.31426	6.82856	8.61646
2	0.112815	11.500	1.0840	16.959	5.982	6.124
3	0.051551	5.255	0.4179	6.19183	4.38816	3.0121
4	0.847 584	86.400	0.823	12.608	5.481	4.958

Table (11) Experimental and theoretical results of model 5, 6, 7 and 8

Test No.	Experimental results		Theoretical results			
	$Q_{cr,}$ kN	$Q_{cr,}$ kg	ρ	μ	s	sc
5	18.474	1883	0.1841102	1.431531	1.862895	0.292099
6	10.607	1081	0.1057058	5.460222	1.175894	-0.394902
7	4.735	482	0.047188	17.19175	0.922095	-0.648701
8	1.581	152	0.015755	59.4687	0.826299	–0.744498

The buckling load is obtained for the first model divided into two, three and four non-prismatic elements. Hereafter, the stability functions and non-dimensional axial force parameter for each element are obtained and given in Table (12).

Table (4) Stability Functions for model 1 divided into 1, 2, 3 and 4 tapered members

Element	L m	Q kN	Stability Function	Values of Stability Functions			
				First	Second	Third	Fourth
One	0.3	0.00394	S_1	27.3143	-	-	-
			SC	8.61646	-	-	-
			S_2	6.82856	-	-	-
			ρ_2	1.66590	-	-	-
Two	0.15	0.00394	S_1	12.6399	9.3238336	-	-
			SC	4.63927	3.5783314	-	-
			S_2	5.61763	5.2446338	-	-
			ρ_2	0.416475	0.0822700	-	-
Three	0.01	0.00394	S_1	9.139763	7.7054150	6.8649449	-
			SC	3.613771	3.1403804	2.8841442	-
			S_2	5.141116	4.9314656	4.7673228	-
			ρ_2	0.185100	0.0585664	0.0239888	-
Four	0.075	0.00394	S_1	7.629459	6.8353036	6.3118599	5.9462715
			SC	3.155855	2.8903702	2.7254391	2.6124706
			S_2	4.882853	4.7467386	4.6372774	4.5526189
			ρ_2	0.104119	0.0426468	0.0205665	0.0111013

The experimental data for models from 1 to 8 are identical with theoretical results that used modified stability functions by including shear effect. From the previous results, the elastic critical load of models 5, 6, 7 and 8 are compared with others, as shown in Figure (10) to represent the effect of increasing the number of batten lacings on elastic critical load in a prismatic member.

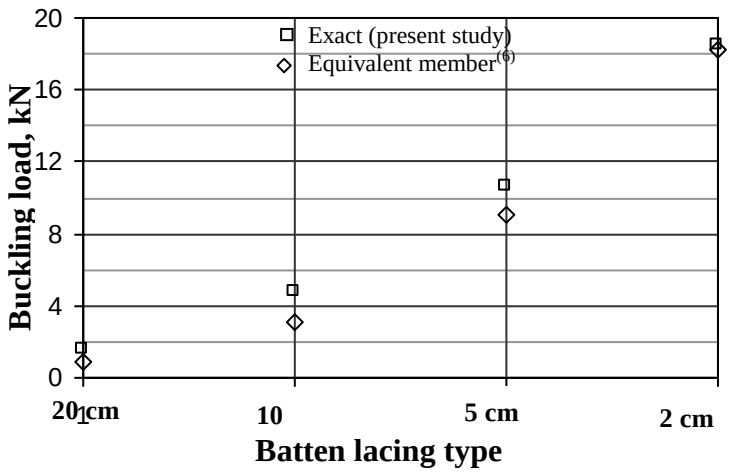


Fig (10): Buckling load on prismatic member for different batten lacings

It is found that the prismatic member having 2cm-batten lacing buckled after the three other models which have less number of battens lacing. On the other hand, the prismatic member having 20cm batten lacings buckled before the three other models which have a larger number of battens lacings. This means that, when the number of batten lacing increased, the buckling load increased. The other type of comparison is presented in Table (13) for the elastic critical load of the present study and the equivalent member [(6)]. It is found that the elastic critical load, which is obtained from the present study, is more than that obtained from the equivalent member. Their values are close to each other. The error percent reduces when the number of batten lacings is increased.

Table (13): Buckling load and displacement comparisons

Test	No. of batten lacings	Buckling load, kN		Ratio*	Error** %	Displacement at critical load, m	
		Present study	Equivalent member			Present study	Equivalent member
5	20	18.474	18.2631	0.9886	1.14	0.2163	0.2080
6	8	10.6066	9.03929	0.8522	14.78	0.1352	0.1315
7	4	4.735	3.10014	0.6547	34.52	0.0617	0.0512
8	2	1.5809	0.85348	0.5399	46.01	0.0272	0.0191

* The ratio between the buckling load of equivalent member to that of the present study.
** Error = Deference between the buckling load of the present study and equivalent member divided by the buckling load of the present study.

From the above experimental works and theoretical results, the stability functions are very close.

Conclusions

1. The modified stability functions including the effect of bending and shear are compared with the stability functions including the effect of bending only for the same properties of non-prismatic beam-columns under different axial forces and cross-sectional areas. These conditions give different ratios of stability functions, which include and exclude shear effect. The effect of shear is summarized in Table (14) which shows: -

Table (14): Ratios between stability functions excluding and including the effect of shear

μ_2 \ ρ_2		0.00	2.00
0.0001	S_1	0.999747	0.999813
	SC	0.999600	0.999888
	S_2	0.999838	0.999998
0.0400	S_1	0.908997	0.928652
	SC	0.854400	0.958576
	S_2	0.941763	0.999844

a. The effect decreases the stability functions with increasing of non-dimensional axial force parameters in the compression range.
b. The effect increases the stability function with increasing of value of shear parameter.
c. The minimum effect is at maximum axial force and minimum shear parameter (0.0001).
d. The maximum effect is at zero axial force and maximum shear parameter (0.04).

e. The effect of shear parameter exceeds the effect of non-dimensional axial force parameter.

2. In batten laced members, the shear flexibility parameter is decreased with increasing the number of batten lacings between two main columns. The limit of buckling load is increased by reducing the additional deformation due to shear strain as given by which summarizes the buckling load data for models 5, 6, 7 and 8.

3. .It is found that the prismatic member having 2cm batten lacing buckled after the three other models which have less number of battens lacings. On the other hand, the prismatic member having 20cm batten lacings buckled before the three other models which have a larger number of battens lacings. This means that, when the number of batten lacing is increased, the buckling load is increased.

4. It is found that the elastic critical load, which is obtained from the present study, is larger than the elastic critical load obtained from the equivalent prismatic member given by Timoshenko formula [(6)] .Their values are close to each other and the error percent is reduced when the number of batten lacings is increased.

REFERENCES:

1. Al-Quraishi, H. A. A., “Large Displacement Elastic Stability Analysis of Plane Frames Including Shear Effect”, M.Sc. Thesis, University of Technology, Iraq, 1999.
2. Al-Fadul, M. A., “Stability Functions for Non-Prismatic Members Including Shear Effect”, M. Sc. Thesis, University of Kufa, Iraq, 2005
3. Al-Sarraf, S. Z., “Shear Effect on the Elastic Stability of Frames”, The Structural Engineer, June 1986, pp. 43-47.
4. Lin, F.J., Glauser, E.C., and Johnston, B.G., “Behaviour of Laced and Battened Structural Members”, Journal of the Structural Division, ASCE, Vol.96, No. ST7, July 1970, pp. 1377-1401.
5. Lindgren, S., “Shear Flexibility”, Journal of the Structural Division, ASCE, Vol.105 No. ST10, October 1979,Technical Note 2117-2121.
6. Timoshenko, S. and Gere, J.M., “Theory of Elastic Stability”, 2nd edition, New York, McGraw Hill Book Co., Inc., 1961.
7. Al-Sarraf, S.Z., “Discussion of Frames of Solid Bars of Varying Cross Sections”, Journal of the Structural Division, ASCE 81, No. ST1, February 1985, pp 318.
8. Al-Sarraf S.Z., “Elastic Stability of Frameworks”, Ph.D. Thesis presented to the University of Liverpool at Liverpool, England, July, 1964.
9. Al-Sarraf, S. Z., “Elastic Instability of Frames with Uniformly Tapered Members”, The Structural Engineer, March 1979, pp. 18-24.
10. Yousif, W. V., “Modified Stability Functions with Shear Effects for Non-Prismatic Members in Steel Plane Frames and Members Inside Soils”, Ph.D. Thesis, Department of Civil Engineering, University of Baghdad, April 2006.

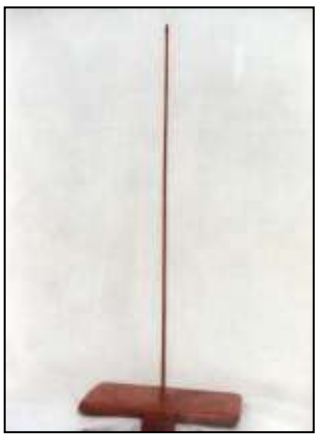
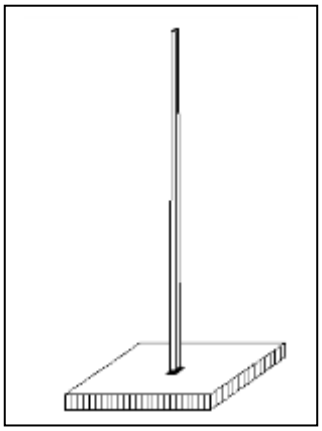
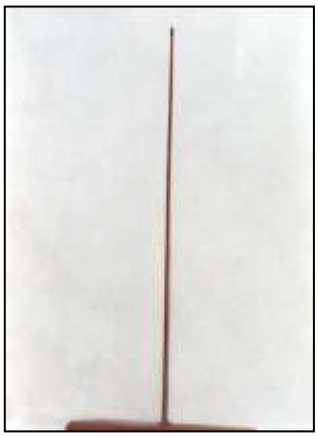
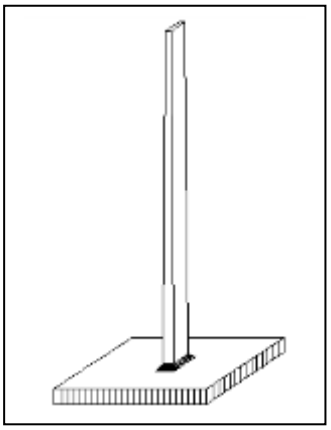
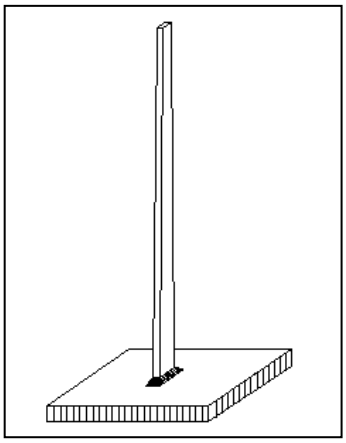
Appendix

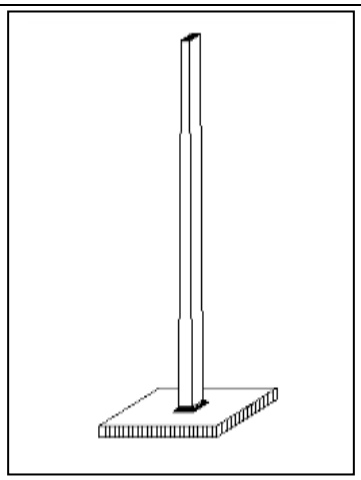
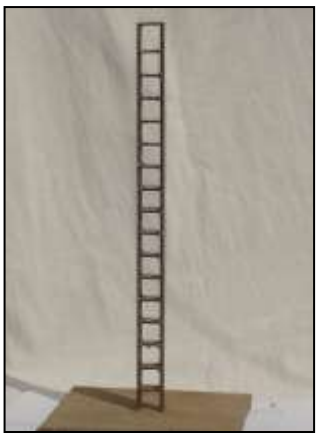
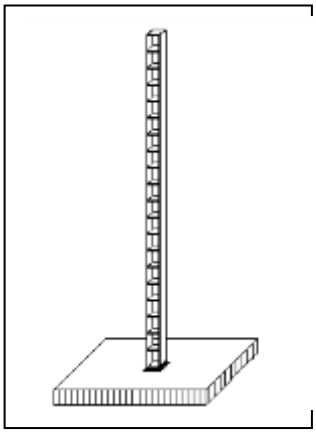
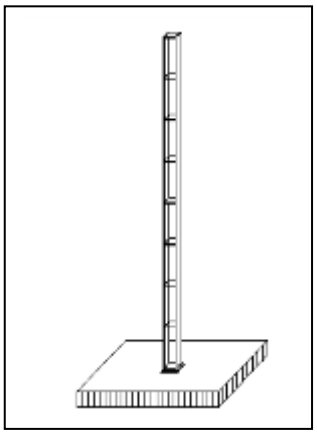
m=3	$S_1 = \frac{Z_1}{EI_2}\frac{K_a(D_1QL + B_1QL + 1) + \mu_2.\rho_2}{Q}$
	$SC = \frac{Z_1}{EI_2}\frac{K_b(C_1QL + A_2QL + 1) + \mu_2.\rho_2}{Q.u}$
	$S_2 = \frac{Z_1}{EI_2}\frac{K_b.u(A_1QL + A_2QL + 1) + \mu_2.\rho_2}{Q.u}$
m=1	$S_1 = \frac{Z_1}{EI_2}\frac{K_a(D_1QL + B_1QL + 1) + \mu_2.\rho_2}{Q}$
	$SC = \frac{Z_1}{EI_2}\frac{K_b(C_1QL + A_2QL + 1) + \mu_2.\rho_2}{Q.u}$
	$S_2 = \frac{Z_1}{EI_2}\frac{K_b.u(A_1QL + A_2QL + 1) + \mu_2.\rho_2}{Q.u}$
m=2	$S_1 = \frac{QLb}{PEI_2}\left[L\beta\cos\psi - 0.5(b+a)\sin\psi\right]$
	$SC = \frac{QL\,a\,b}{PEI_2}\left(\sin\psi - \frac{\beta L}{a^{0.5}\,b^{0.5}}\right)$
	$S_2 = \frac{QLa}{PEI_2}\left(\beta L\cos\psi - 0.5(a+b)\sin\psi\right)$

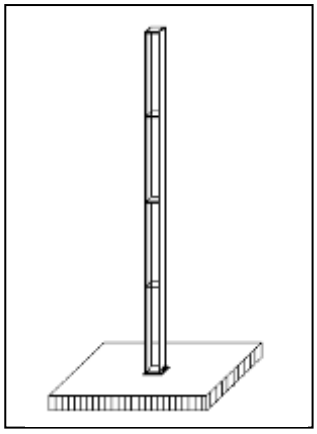
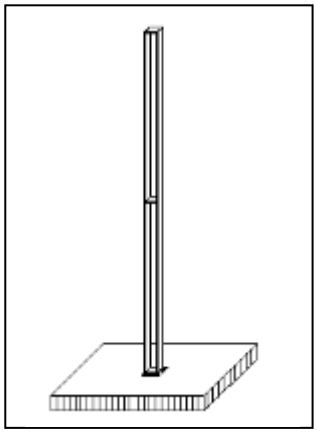
m=4	$Z_1 = L/\left[K_a\frac{\mu_2.\rho_2}{Q.u^2}(D_1 - B_1) - K_bK_aL\left(A_2(B_1 - D_1) + C_1B_1 + B_2(C_1 - A_1) + \frac{C_1 - D_1}{QL} + \frac{B_1 - A_1}{QL} - A_1D_1\right) - K_b\frac{\mu_2.\rho_2}{Q}(C_1 - A_1)\right]$
m=3	$Z_1 = L/\left[K_a\frac{\mu_2.\rho_2}{Q.u}(D_1 - B_1) - K_bK_aL\left(A_2(B_1 - D_1) + C_1B_1 + B_2(C_1 - A_1) + \frac{C_1 - D_1}{QL} + \frac{B_1 - A_1}{QL} - A_1D_1\right) - K_b\frac{\mu_2.\rho_2}{Q}(C_1 - A_1)\right]$
m=1	$Z_1 = L/\left[K_a\frac{\mu_2.\rho_2}{Q.u}(D_1 - B_1) - K_bK_aL\left(A_2(B_1 - D_1) + C_1B_1 + B_2(C_1 - A_1) + \frac{C_1 - D_1}{QL} + \frac{B_1 - A_1}{QL} - A_1D_1\right) - K_b\frac{\mu_2.\rho_2}{Q}(C_1 - A_1)\right]$

m=2	$A = -\frac{M_1 + M_2\,u^{0.5}\cos\psi}{Q\,u^{0.5}\sin\psi}$
	$B = M_2 / Q$

	$\beta = \sqrt{\frac{\bar{\rho}_2 \pi^2}{(u-1)^2} - 0.25}$
	$\psi = \beta \ln\left(\frac{b}{a}\right)$
	$P = L(\beta^2 - 0.25)\sin\psi + (a+b)\beta\cos\psi - 2\beta a^{0.5} b^{0.5}$

Model 1		
Model 2		
Model 3		

Model 4		
Model 5		
Model 6		

Model 7		
Model 8		

List of symbols

Symbol	Symbol Definition
A_v	Effective shear area for prismatic members
A_b	Cross sectional area for batten
A_{v2}	Effective shear area for tapered members at end 2
E	Modulus of elasticity
I_2	Moment of inertia for tapered member at end 2
I(x)	Moment of inertia at distance x from the origin O for tapered member
I_c	Moment of inertia for vertical plate
I_h	Length of the batten center to center of vertical plate
I_v	Center to center vertical distance between two battens lacing
M_1, M_2	Applied moment at end 1 and 2 respectively
M(x)	Bending moment at distance x from the origin O
Q	Constant axial force
Q_E	Euler load ($\pi^2 EI / L^2$)
S_1	Flexural stiffness factor for tapered member at end 1
S_2	Flexural stiffness factor for tapered member at end 2
SC	Flexural carry-over factor for tapered member

V	Shear force
a	Distance of end 2 from the origin O for tapered member
b	Distance of end 1 from the origin O for tapered member
d_1, d_2	Depth of tapered member at end 1 and 2 respectivily
d(x)	Depth of tapered member at distance x fro origin O
f	Frequency in cycle /sec
g	Acceleration due to gravity
k	Stiffness of structure
m	Shape factor
n	Numerical factor equal to 1.2 in the case of a rectangular cross section
u	Tapering ratio
W	Summation of external load on frame
y	Lateral deflection at distance x along the member
μ	Shear flexibility parameter
θ_1, θ_2	End rotations at end 1 and 2
ρ_2	Non-dimensional axial force parameters for tapered member

الكفاءة الإستخدامية للمناطق المكشوفة والخضراء لكلية الزراعة

نجوى عبيد عجمي
كلية الزراعة – جامعة بغداد

المستخلص :-

ان الهدف الاساس من عملية تخطيط وتصميم الفضاءات الخارجية والحدائق هو خلق بيئة خارجية ملائمة لاستخدام الانسان تتوازن فيها الوظيفة مع المتعة مع الاخذ بنظر الاعتبار ظروف الموقع وخصوصيته ومميزاته وامكاناته الاقتصادية . يهدف البحث الى تحديد اهم السلبيات والمشاكل التي تعاني منها المساحات المكشوفة والحدائق في كلية الزراعة مع دراسة تجارب الدول العربية لبلدان مشابهة لظروف العراق البيئية واستباط القاعدة المعلوماتية التي يمكن ان تكون ورقة عمل لتحسين وتطوير وديمومة هذه المساحات المفتوحة والخضراء وصيانتها وتشغيلها بالشكل الذي يحقق الغرض من انشائها . تضمنت الدراسة جانب نظري اشتمل على توضيح اسس تخطيط وتصميم وصيانة الحدائق والعوامل المؤثرة عليها وعرض لتجارب دول عربية ، وجانب عملي تضمن تحديد المشاكل والسلبيات التي تعاني منها المساحات المكشوفة والحدائق في كلية الزراعة واستمارة الاستبيان ومقترحات التطوير .

EFFICIENT EMPLOYABILITY OF OPEN GREEN AREA IN THE AGRICULTURAL COLLEGE – UNIVERSITY OF BAGHDAD

ABSTRACT

The main goal of planning and designing for external open spaces and gardens is to create a comfortable external environment for human use. In this environment, there should be a balance between the function and the aesthetic, taking into account the location's circumstances, specialty, characteristic and economic potentials. The research is aimed at limiting the disadvantages and the problems of the opening areas and gardens in the agriculture college supported by study of the similar Arabic countries experience that have the same Iraq circumstances environment.Also building a data base that may be a work sheet to improve, develop and ever-living these opening green areas and maintain and act it in the form that achieves the purpose behind it's establishment.

The study includes a theorem side which includes illustration of foundations of planning , designing and maintain the gardens and the external infected factors and show the Arabic countries experience, and application side that includes limiting the problems and disadvantages of the opening areas and gardens in the agriculture college and the information form as well as the development suggestions .

Keywords: Open and Green Areas, Efficient Employment, Agriculture College .

المقدمة

إن هدف التخطيط هومحاولة وضع خطة ستراتيجية ضمن اسس علمية لغرض استغلال الموارد البشرية والطبيعية والمادية للمواقع الاكاديمية والبحثية وعلى هذا الاساس فان العملية التخطيطية تتطلب توافر عاملين الاول الهدف والثاني المحددات والقيود التي تحدد التنفيذ وعلى المخطط التوفيق بين الاهداف والمحددات بشكل عقلاني وموضوعي . كما انه لا يمكن فصل العملية التخطيطية عن العملية التصميمية فهما عمليتان متلازمتان للعلاقة الوثيقة بينهما . ان انشاء طابع محلي يساهم في أغناء الشخصية المعمارية ويعكس المحتوى الحضاري للمدينة ضمن فضاءات الكلية ليبقي في ذاكرة الطالب دوماً وكله فخر واعجاب بروعة التصاميم وكفاءة الاداء للفضاءات المفتوحة والخضراء وجمالها وهي غاية يسمو اليها المخطط والمصمم معاً . ان تنسيق وتجميل الحدائق يمثل احد اركان التطوير والتنمية العمرانية ، فدراسة تصاميم المناطق الخضراء والحدائق والفراغات وممرات المشاة واماكن الجلوس واسس اختيار النباتات الملائمة للبيئة كل ذلك يساعد على توفير بيئة ملائمة وخلق جو دراسي مريح للطالب ومن هذا المنطلق جاءت أهمية دراسة الفضاءات المفتوحة لكلية الزراعة في جامعة بغداد منطقة ابو غريب وما لها من خصوصية تميزها عن باقي الكليات حيث المساحات المكشوفة والحدائق لها وظيفة تعليمية وبحثية الى جانب الوظيفة التخطيطية والبيئية والجمالية والترفيهية ، ولوضع الأطر التصميمية والتخطيطية لتطويرها تضمن البحث مرحلتين : المرحلة الاولى / تشمل الجانب النظري الذي تطرق الى أسس تخطيط وتصميم وصيانة الحدائق والعوامل المؤثرة عليها ، والمرحلة الثانية / تشمل الجانب العملي والمتضمن واقع حال حدائق كلية الزراعة في جامعة بغداد وتحديد المشاكل التي تعاني منها والتي توضحت عبر استمارة الاستبيان وما ترتب عليها من مقترحات تطويرية .

- أسس تخطيط وتصميم وصيانة الحدائق

للمناطق الخضراء قيمة كبيرة باشجارها وشجيراتها وازهارها ومسطحاتها الخضراء فهي تمثل واجهة وشخصية المدينة ، ولكي نصل الى الحالة المثلى لابد لنا من معرفة اساسيات التصميم والتخطيط للحدائق فالتصميم بمعناه الشامل هو تنظيم الاجزاء البسيطة في صورة مركبة وبطريقة فنية للوصول الى تنظيم وبالتالى تنسيق جيد فلكل حديقة محور رأسي طولي ومحور او اكثر ثانوي او عرضي عامودي على الرئيسي ولكل محور بداية ونهاية يبدأ بعلامات دالة او عناصر جذب كنافورة او نصب ولابد من تحقيق الوحدة والترابط التي بين اجزاء الحديقة بفكر تصميمي موحد يجعل من جميع اجزاء ومكونات الحديقة تتسم بالترابط والتناسب والتوازن فهناك وحدة التصميم لجميع عناصر الحديقة سواء بالالوان او بالمواد المستخدمة من رصف المماشي ، الاكشاك ، مظلات ومقاعد الجلوس، الحاويات...الخ ، وبصفة عامة فان أي حديقة تتألف من العناصر الرئيسية التالية : www.momra.gov.sa/Specs/guid0021.asp

اولاً - العناصر النباتية : تعتبر النباتات العناصر الاساسية التي تتكون منها الحديقة فهي ذات وظيفة تخطيطية حيث تعمل على تحديد وتقسيم المساحات والفصل بين المرافق المختلفة وتحديد المماشي ومسارات الحركة وتكوين اسوار نباتية لحجب المناظر غير الرغوب فيها ، كما أن لها استخدام بيئي في مكافحة التلوث وامتصاص الغازات غير المرغوب فيها وتقليل الضوضاء ومقاومة التلوث البيئي وتلطيف درجة الحرارة ولنشر الظل خاصة في المناطق الصحراوية وكسر حدة الرياح ، ولها استخدام جمالي حيث تشكل الاشجار والنباتات الاخرى العنصر الاساسي لجمال المدن وتنسيق المواقع والحدائق وعنصر جذب بجمال اشكالها والوانها وقابليتها على القص والتشكيل www.momra.gov.sa/specs/guid0014.asp لتعطي احساس للزائر بأن هناك ترتيب في استخدام النباتات و الالوان بشكل يضيف الجمال ويبعث المتعة في نفس المتلقي ، وكما موضح ذلك في (شكل -١) وينبغي ان تكون النباتات المختارة تؤدي الدور المطلوب منها على اكمل وجه وللمملكة السعودية تجربة رائدة في مجال الاهتمام بزيادة المساحات الخضراء وحمايتها و التحديد الصحيح لمواقعها وتنظيم استغلالها وفي مدينة ينبع تم عمل مسح شامل لجميع اصناف النباتات المحلية والمستوردة لتوظيفها في انظمة التشجير بعد اجراء التجارب والابحاث وانتخاب الانواع التي تتحمل الظروف المناخية المحلية مع قلة احتياجها لمياه الري وعمليات الصيانة ومدى تناسبها لغايات واهداف التشجير حيث تم تطبيق المعايير الاساسية للتشجير في اختيار العناصر

النباتية بناء على دراسة الموقع والغرض منها وتجهيز التربة وحساب الاحتياجات المائية للنباتات واستخدام قائمة النباتات المعتمدة من قبل الجهات المختصة لكل مدينة وتزويد المناطق المشجرة بنظام الري (الالي) المبرمج بشكل دائم وثابت وعندما تستخدم الاشجار كاسوار يجب ان يتالف السور من اشجار وشجيرات دائمة الخضرة عريضة الاوراق ومتقاربة المسافات ، وتزرع الاشجار والشجيرات كنماذج فردية او في مجاميع حسب استخداماتها المختلفة لتكسب المكان منظراً جميلاً كما تزرع النباتات العشبية الحولية والمعمرة لالوان ازهارها المتعددة واهميتها في عمليات التنسيق وتزرع احواض الزهور في خليط لا يتعدى اكثر من ثلاثة انواع من الازهار مع مراعاة ترتيب الالوان وتوزيعها بحيث تعطى تكويناً متوازناً خلال فصل النمو. وضمن العناصر التكوينية الاتية :

www.urar.org.sa/ibda/mahawer3 - 4.html

شكل (١) جوانب من الحدائق والاستراحات في السعودية التي توضح العناية بزراعة النباتات وتنسيقها لتكسب المكان منظراً جميلاً

المصدر : www.arriyadh.com/Tourism/LeftBar/Gardens

ثانياً - العناصر البنائية : وتشمل ممرات المشاة ،المقاعد واماكن الجلوس ، المظلات ، المجسمات البنائية ، الاحواض البنائية ، حاويات النفايات . يوجد في الحديقة عدد من الممرات التي تربط مداخل الحديقة وأجزائها وتوصل الاماكن المختلفة فيها وعند انشاء هذه الممرات يجب ان يراعى طراز الحديقة المستعمل ويجب الاهتمام بالنواحي البصرية على جميع محاور وممرات المشاة وخاصة التي في مستوى النظر لاعطاء متتابعات بصرية متنوعة وممتعة http://civilawy.jeeran,com/lectures2.htm كما ان المواد المستخدمة في اكساء الارضيات لها معايير تخطيطية ووظيفية وجمالية يجب ان تؤخذ في الاعتبار ولكل مادة من مواد الاكساء تأثير مختلف عن الاخر من خلال التنوع في المقاسات والاشكال والملمس والالوان وقوة التحمل والمقاومة للعوامل الجوية والحاجة الى الصيانة المستمر كالخرسانة ، الحجر ، البلوك ،الاسفلت ويختلف عرض المماشي ونوع المواد المستخدمة في ارضيتها حسب نوع الحديقة ومساحتها وحسب طرازها . كما يراعى توفر اماكن الجلوس ويعمل على رصف الطرق المؤدية اليها وتجنب وضعها على المسحطات الخضراء لرطوبتها المستمرة بل يخصص منطقة للجلوس يوضع بها رمل او ترصف بالبلاط كما يتوقف تصميمها على طراز الحديقة والغرض الذي تنشا من اجله كمكان منعزل يشعر فيه الانسان بهدوء الطبيعة او الاستراحة او كمكان لتناول الطعام مع وجود بعض المقاعد والطاولات البنائية كما ان موقع اماكن الجلوس ونوعية المقاعد المستعملة فيها لها اهمية كبيرة في دراسة النواحي الوظيفية والجمالية وعموما فان اماكن الجلوس يجب الا تعترض انسيابية الحركة في الممرات

الرئيسية والساحات ويمكن دمج اماكن الجلوس في التكوين مع احواض الزرع بحيث تكون هذه الاماكن مواجهة لمحاور حركة المشاة ، كما يمكن استخدام حواجز ومحيطات حوض الزرع كاماكن للجلوس وفي هذه الحالة يراعى ان تكون بارتفاعات مناسبة ومريحة . ومن العناصر البنائية المهمة وبالاخص في المناطق الحارة المظلات حيث ان تظليل الطرق والممرات واماكن الجلوس ذو قيمة جمالية ومهمة في الاماكن المشمسة بهدف تهيئة الراحة ، وتعتبر الاقواس من المنشأت المعمارية المستخدمة كدعامات للمتسلقات وتجميل البوابات والمداخل واذا وضعت فوق الطرق الطويلة فانها تكسر من حدة هذا الطول وما يبعثه من ملل .

ثالثاً ـ العناصر المائية : يعتبر توفر الماء من اكثر العوامل أهمية لنمو النباتات ولابد من تأمين الماء لضمان ديمومة النباتات ،ويعتبر الري أحد الفنون القديمة قدم الحضارة نفسها ، وقد حدث هذا ، في مصر والصين وباكستان واليابان وبلاد ما بين النهرين (موقع الدراسة الحالية) وغيرها ، والاداره المتكاملة للري هو تطبيق الأسس والتقنيات المناسبة لتحقيق الاستثمار الأمثل للموارد المائية المتاحة وترشيد استخدامها . ولغرض تحقيق هذا الهدف يجب إجراء الدراسات الخاصة بالاحتياجات المائية للمحصول وتحديد جدول للري (تحديد مواعيد وكميات مياه الري للنباتات) وهندسة الري- هي العلم الذي يهتم بتزويد المساحات الزراعية بالمياه اللازمة للاستخدامات الزراعية بطريقة محسوبة بدقة على أساس المناخ و الطبوغرافيا وطبيعة التربة .

www.uaeagri.ae/home.aspx?ctname=ArticleDetails.ascx&rid=76

وطريقة الري المثلى هي التي تمد الأرض بالرطوبة لنمو النبات دون فاقد في المياه أو التربة ، وتؤمن المحصول ضد فترات الجفاف القصيرة ، وتغسل الأملاح الموجودة في القطاع الأرضي لتصبح دون الحد الحرج للحصول على أكبر وأجود محصول ، مع كفاءة استخدام المياه والتميز في العائد الاقتصادي من وحدة الماء. وعموماً من المعروف تماماً أنه لكي تروي لا بد من استخدام التكنولوجيا (بدءاً من وسائل الري البدائية إلى أحدث نظام للري) لتنظيم إمداد النباتات باحتياجاتها المائية .

وهناك عوامل يجب مراعاتها لاختيار نظام الري حيث يعتبر اختيار طريقة الري الأكثر ملائمة جزءاً هاماً في تخطيط مشاريع الري و لكل طريقة ري مميزاتها الخاصة ، ولكل تصميم مشروع ري مشاكل وصعوبات ، وعموماً يجب أن يراعي في أي نظام ري المعايير الأساسية التالية :

- تلبية الاحتياجات المائية المناسبة للنباتات خلال مراحل النمو المختلفة مع اعتبار المناخ.
- العدالة في توزيع مياه الري توزيعاً منتظماً وصولاً الى منطقة الجذور.
- تقليل عمليات التعرية في سطح الأرض المروية إلى أقل قدر ممكن.
- إنجاز عملية الري بأقل عدد من العمالة اللازمة.
- أن تتمشى طريقة نظام الري مع التغييرات في التربة وطبوغرافية السطح.
- أن يكون لنظام الري مرونة كافية في تطبيق ميكنة العمل المزرعي بالمنطقة المروية
- مراعاة نظام الري كجدوى اقتصادية للمشروع.

ولتوفير معظم المعايير المذكورة أنفاً يجب الإلمام بجميع تفاصيل العوامل المحددة للإنتاج الزراعي حتى يتم التخطيط السليم لمشروع ري يناسب كل الظروف ويؤدي إلى أفضل عائد بأقل تكلفة ممكنة. أما الموارد التي يجب معرفتها قبل تحديد واختيار طريقة الري فتشمل: الموارد المائية (مصدر المياه، كمية المياه، نوعية المياه)، التربة، نوع المحصول، رأس المال والعمالة المتوفرة، الطاقة، نوع المكننة المستعمل، نظام الصرف القائم.

www.uaeagri.ae/home.aspx?ctname=ArticleDetails.ascx&rid=76

ان الري بالرش الرذاذي والتنقيط هو أحد أنظمة الري الحديثة و التي تستخدم لرى المناطق الصحراوية ذات الأرض الرملية و التي لا تستطيع الإحتفاظ بالماء لمدة طويلة، حيث إن تطبيق نظام الري بالغمر يسبب فقد الكثير من الماء مما ينتج عنه إهدار مياه الري، في حين الري بالرش يوفر الماء حيث إن متوسط كفاءة الري لهذا النظام هي.% 75 ، هي مناسبة أيضا في ري الأراضي التي تروى بالرفع من الآبار الارتوازية. و في هذه الطريقة يلزم دفع المياه من مصادرها المختلفة بإستخدام موتورات مناسبة القوة في شبكة مواسير من الحديد المجلفن أو البلاستيك (P.V.C) تتناقص أقطار هذه المواسير تدريجيا

كلما تباعدت عن مصادر المياه .وفي مدينة الكويت تجربة واسعة في استخدام نظام الري بواسطة الرشاشات الصغيرة وتعتبر من افضل الطرق المعتمد استخدامها في ري المسطحات الخضراء .
حيث تم الاعتماد على القطاع الخاص من خلال الشركات المؤهلة في تنفيذ مجمل عقود ومشاريع الزراعة التجميلية (سواء عقود التنفيذ او عقود الصيانة) وتم الاعتماد في ري معظم مواقع الزراعات التجميلية على أنظمة الري الحديثة ، وذاتية التشغيل التي تحد كثيرا من الاستهلاك المائي (التنقيط - الرش) باستخدام نوعيات مختلفة من المياه (المجاري المعالجة - الصليبية - العذبة) وتتوجه الهيئة من خلال مشروعاتها المتعاقبة نحو تعميم نظم الري الحديثة على كافة المواقع والمشاريع واحلال مياه المجاري المعالجة بديلا عن المياه الصليبية والعذبة في ري مجمل الزراعات التجميلية والحرجية،كما يتم الاستفادة من مياه المجاري المعالجة كنوعية جيدة تسهم في تحسين خواص وخصوبة التربة ونمو وازدهار جميع الزراعات.ان نظام الري المستخدم في شارع الخليج العربي ومواقع اخرى في مدينة الكويت بواسطة الرشاشات الصغيرة افضل نظم الري المعتمد استخدامها في ري المسطحات الخضراء والمواقع المشابهة بمختلف انحاء العالم ولاسيما باستعمال شبكة ذاتية التشغيل تسمح بالتحكم في مواعيد الري والتوقيت المناسب للوفاء بالاحتياجات المائية الفعلية دون هدر يمثل اكثر من ٤٠ في المئة في حالة الري التقليدي ويستخدم هذا النوع من الري في كثير من مناطقنا القريبة على مثال موقع مدينة الثرثار الجديد ومقترباته .
كما يحقق نظام الري بالرشاشات توزيعا متماثلا للمياه على مجمل المسطح الاخضر يسمح بالنمو المتجانس لجميع اجزاء المسطح بما يعكس قيمة فنية وتجميلية يلحظها الجميع بسهولة.وعليه، فان خروج بعض قطرات المياه اثناء الرش امر يصعب التحكم فيه بشكل مطلق نظرا للتباين الحادث من وقت لاخر في سرعة الرياح واتجاهاتها ولكن تولي الهيئة هذا الامر اهتماما بالغا اذ يقوم المختصون بضبط ومعايرة فتحات الرشاشات وتواقيت وزمن الري بما يسمح بالتغلب على هذه المشكلة وذلك من آن لاخر وفقا للمناخ السائد والاحتياجات المائية التي تختلف من وقت الى اخر.
www.uaeagri.ae/home.aspx?ctname=NewsDetails.ascx&news_
ان طريقة الري بالرش لها فوائد من حيث تقليل كمية المياه المستخدمة في الري ورفع كفاءته ، الا ان استخدام هذه الطريقة خصوصاً بمياه ذات تركيز مرتفع من الاملاح قد يسبب تراكم الاملاح على اسطح اوراق النباتات خصوصا في فصل الصيف نتيجة لارتفاع درجة الحرارة .
اما ا لري بالتنقيط هو ذلك النظام الذي يتم فيه إضافة المياه للتربة وفي صورة قطيرات صغيرة إلي منطقة الجذور. وينفرد الري بالتنقيط عن غيره بأنه يقوم بترطيب جزء من التربة فقط وتبقي الأجزاء الأخري جافة طوال الموسم وينتج عن هذا الترطيب الجزئي فوائد عديدة ومشاكل قليلة. ويتم إضافة المياه في منطقة جذور النباتات فقط أما المنطقة التي ليس بها جذور فلا يضاف لها مياه وبالتالي التوفير في كميات المياه المضافة ويقلل من مشكلة ملوحة التربة في منطقة الجذور(شكل -٢) يوضح نموذج لشبكة ري بالتنقيط .
يمكن إستخدام مياه ذات ملوحة عالية نسبياً والتي لا يمكن إستخدامها مع الري بالرش أو الري بالغمر، سهولة تحويل النظام ليعمل أوتوماتيكياً بالكامل ،.يستخدم تقريباً نصف إلي ثلثي كمية المياه اللازمة للري بالرش .التكاليف الإستثمارية في البداية عالية. www.zira3a.net/forum/showthread.php?t=708 -
وتعتبر طريقة الري بالتنقيط من أكفأ طرق الري الحديثة الا أنه عند ارتفاع ملوحية مياه الري وارتفاع درجة الجو في فصل الصيف فانه يجب التاكد من ان فتحات المنقطات واسعة لكي تعطي تدفقاً عالياً وبالتالي منع تراكم الاملاح حول جذور النباتات بعد تبخر الماء من التربة لارتفاع درجة الحرارة .
اما طريقة الري بالبلرز (النبع) وهي تحديث وتحسين الري بالتنقيط حيث لوحظ ان الفتحات التي يخرج منها الماء في الري بالتنقيط كثيراً ما تغلق بالاملاح او بحبيبات التربة فاستغنى عن الصمامات في هذه الفتحات باستعمال أنبونتين واحدة داخل الاخرى يخرج ماء الري منها نتيجة لفروقات الضغط .
وينصح باستخدام الطرق الثلاث الاخيرة لتوفيرها في كميات المياه المستهلكة ولسهولة استعمالها .

شكل -٢ يوضح نموذج لشبكات الري بالتنقيط

المصدر : البطراوي ، احمد ، شبكات الري بالتنقيط .www.zira3a.net/forum/showthread.php?t=708 -

اما الري بالغمر او الري السطحي فيؤدي الى اهدار كمية كبيرة من الماء الذي يفقد معظمه بالتبخير من سطح التربة، وعلى الرغم من ان جزءا منه يعود الى التربة لتغذية مصادر المياه الجوفية السطحية الا ان هذا الجزء يعود محملا بالاملاح وملوثا بالمخصبات الكيماوية الزراعية ومبيدات الاعشاب والآفاق مما يسبب تلويث المياه الجوفية ويجعل تأثيره اكثر ضررا.كما ان طريقة الري المحوري التي تستخدم اسلوب الرش عاليا في الهواء هو أحد الاساليب التي تؤدي الى اهدار كبير في المياه نتيجة التبخر السريع مما يسبب استنزاف المياه الجوفية بشكل غير مبرر خصوصا في مثل مناطقنا الحارة الجافة التي تساعد على سرعة التبخير، واذا ما تم تحويله الى نظام الرش الدقيق منخفض الضغط الذي يوصل مياه الرش الى اقرب نقطة ممكنة من المحاصيل المزروعة عن طريق انابيب مدلاة عموديا من ذراع الرشاش فإن ذلك يرفع من كفاءة استخدام المياه في الري بشكل كبير خاصة لو تم الى جانب ذلك استخدام الطرق الحساسة الرخيصة لتقدير رطوبة التربية وتحديد مواعيد الري المطلوبة مثل طريقة مكعبات الجبس المدفونة فإن ذلك لا يؤدي فقط الى رفع كفاءة استخدام المياه في الري بل يزيد انتاجية المحصول في نفس الوقت لضمان حدوث الري في الوقت المناسب بالكمية المطلوبة فقط.

وفي نفس الوقت فان التوسع في نظام الري بالتنقيط مطلوب مهما كانت تكلفة تنفيذه مرتفعة اذ انها سيتم استرجاعها في صورة توفير مياه الري خلال زمن قصير ، وللمملكة الاردنية الهاشمية تجربة ناجحة في ذلك تم تنفيذها في احد الوديان الكبيرة خلال ربع القرن الاخير ادى فيها استخدام طرق الري بالتنقيط الى استخدام نفس كمية المياه المتاحة في ري مساحة تزيد عشرة اضعاف عن المساحة الاصلية التي كانت تروى بالغمر وارتفعت انتاجية الارض في نفس الوقت ما بين ثلاث الى خمس مرات.

– www.fao.org/docrep/007/ad820a/ad820a03.htm/ شهدت مدينة السعودية تطوراً ملموساً وزيادة واضحة في انتاج المياه ، وتوفير المياه المستعملة في عمليات الري وحظيت مدينة ينبع الصناعية بتطبيق أنظمة الري الحديثة وجعله ضمن شبكة التجهيزات الاساسية فاصبحت نموذجاً مميزاً يحتذى به بين المدن التي تواكب تطورات العصر . ان مدينة ينبع تقع على ساحل البحر الاحمر وفي منطقة صحراوية تعاني من ندرة المياه العذبة ومياه الري فتم ايجاد البديل بانشاء محطة لتحلية المياه المالحة (مياه البحر) ومحطة اخرى لمعالجة مياة الصرف الصحي ، www.urar.org.sa/ibda/mahawer3-4.htl

والجدير بالذكر ان اخر الابحاث التي اجريت في قسم وقاية النباتات في كلية الزراعة جامعة الملك سعود

والتي استمرت لثلاث سنوات ، أظهرت أنه لا ضرر من استخدام مياه الصرف الصحي المعالجة (معالجة ثنائية على الاقل ويفضل أن تكون المعالجة ثلاثية) في ري نباتات الزينة وأنها أمنة في محتواها الكيمياوي والميكروبي ، وقد أوصت الدراسة باستخدامها خاصة للحدائق العامة (عز الدين الفراج، ١٩٨٦) حيث أنها خالية من الروائح التي تزعج المواطنين ولا يكون هناك أضراراً وسلبيات من استخدامها . وللماء استخدامات اخرى حيث ترجع أهمية استخدام العناصر المائية والنافورات في الحديقة الى تأثيراتها الجمالية والوظيفية وذلك من خلال شكل التكوينات المائية وجمال مظهرها وحركة الماء الانسيابية وخرير صوته ،بالاضافة الى الدور الهام الذي تقوم به المسطحات المائية في تلطيف درجة حرارة الجو وزيادة الرطوبة النسبية ، وفي حالة المسطحات الخضراء التي تحتاج الى ري مستمر في المناطق الحارة يمكن ان ياخذ نظام الري بالاعتبار في التصميم بحيث يتم اضافته كعنصر جمالي ،وغالبا ما تلعب النافورات دور هام كتكوين جمالي او عنصر تشكيلي ودراسة وضع النافورة في الفراغ بالنسبة لضوء الشمس ودراسة الانعكاسات من او على الماء ولابد لنا من معرفة العوامل المؤثرة على تصميم الحدائق والمتمثلة في : www.momra.gov.sa/Specs/guid0021.asp

أ - الغرض من انشاء الحديقة : يعتبر الغرض من انشاء الحديقة عامل مهم في تحديد التصميم المناسب حيث يختلف تصميم الحديقة العامة عن المنزلية او حدائق المدارس او المستشفيات ولكل منها مواصفات خاصة تلائم الغرض من الانشاء . ونسبة الاستخدام او اسلوب العرض التأثيري .

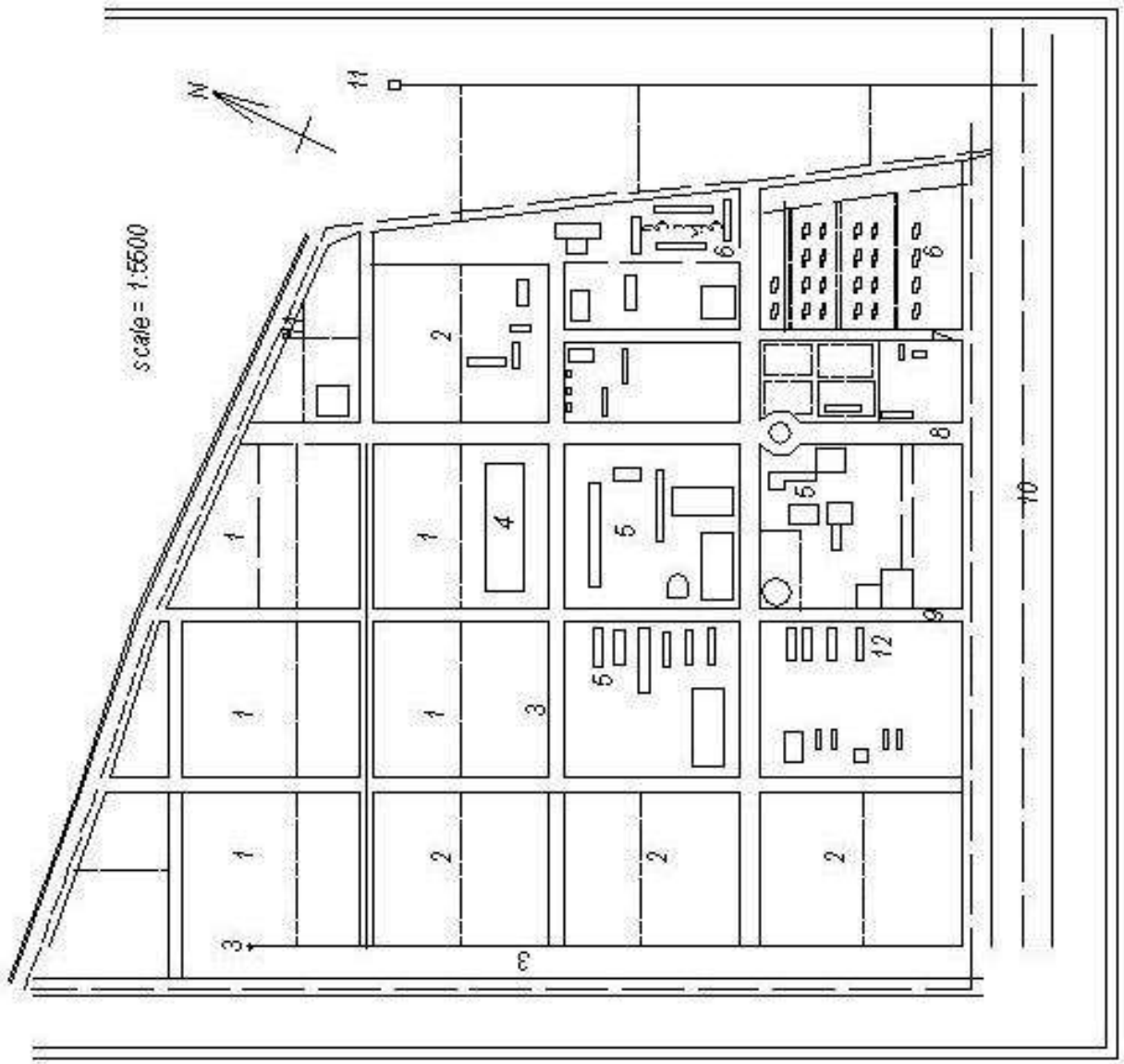
ب - العوامل الطبيعية وتشمل العوامل المناخية وهي من اهم العوامل التي لها تأثير كبير على الحديقة لان الحديقة معرضة بشكل مباشر الى تأثيرات العوامل المناخية من درجة حرارة ورياح ورطوبة وشكل وطبيعة الارض ونوع التربة .

ج - الامكانيات المالية لانشاء الحديقة وصيانتها حيث يتوقف تصميم الحديقة على مدى المقدرة المالية لتغطية المصاريف اللازمة لانشائها واقامة بعض المنشأت البنائية فيها وزراعة النباتات وعمليات الصيانة اللازمة وما تحتاجة من عناية مستمرة .

- وصف لمنطقة الدراسة وتحليلها من قبل الباحث :

تأسست كلية الزراعة في التاسع من شهر تموز من عام ١٩٥٢(مجلة المزارع الحديث ،١٩٥٢) ، تقع كلية الزراعة / جامعة بغداد الى الشمال الغربي من محافظة بغداد ، وهي خارج حدود التصميم الاساسي لمدينة بغداد ، وضمن التصميم الاساسي لناحية ابو غريب ، الحجم الطلابي للكلية على مستوى الدراسات الاولية ٣٨٠٠ طالب وطالبة وعلى مستوى الدراسات العليا ٦٣ طالب ماجستير و٦٦ طالب دكتوراه ، تبعد كلية الزراعة عن مركزبغداد حوالي ٢٢ كيلو متراً ، يحد كلية الزراعة من الشرق معسكرات ابو غريب سابقاً بعد اضافة المعهد الزراعي للكلية ، ويحدها من الغرب ناحية ابو غريب ، ومن الجنوب الشارع العام المؤدي الى محافظة الانبار ، ومن جهة الشمال نهر ابو غريب الذي يمرمحاذياً للحدود الشمالية للكلية ، ومن الخارطة (١) التي توضح استعمالات الارض للكلية ، نلاحظ بان موقع الكلية يشتمل على ثلاثة مداخل جميعها يقع على طريق بغداد / الرمادي كذلك يشمل الموقع على مجموعة من الطرق المتعامدة واماكن لوقوف السيارات وابنية تعليمية تشمل القاعات الدراسية والمختبرات والمكتبة والبيوت الزجاجية والبلاستيكية والمظللات وابنية ترفيهية تشمل النادي وقاعة التاميم والملاعب الرياضية وابنية خدمية للطلاب وابنية لخدمات الصيانة تضم وتشمل الماء والكهرباء والمعمل وابنية للخدمات الخاصة مثل المستوصف فضلا عن الابنية الادارية والجامع كذلك يشمل الموقع على اقسام داخلية للطلاب والطالبات وحي سكني لاساتذة ومنتسبي الكلية ، وشبكة من سواقي الري والمبازل ومحطة مياه المجاري ومنحل وهناك ابنية ضمن موقع الكلية ولكنها لا تخضع لادارتها مثل ابنية المعهد الزراعي الذي الغي وضم الى الكلية والمدرسة الابتدائية . توجد في الكلية ثلاث مداخل هي على التوالي مدخل الحي السكني ومدخل العمادة ثم مدخل الطلبة وجميعها تقع على طريق بغداد / الرمادي . مدخل العمادة هو المدخل الخاص بالاساتذة والموظفين يرتبط بصرياً بساحة السباع وتعتبر نقطة ذات دلالة بصرية المدخل ذو سعة كافية لاستقبال اعداد المنتسبين ولكن غلق احد جوانب المدخل لاغراض امنية يسبب زخم مروري وخاصة في اوقات الذروة ،المدخل يفتقر الى فضاء مسقف محمي مناخياً من اشعة الشمس يمكن الموظفين من انتظار

وسـائط النقـل ،كمـا انـه يفتقـر الـى مخطـط توضـيحي للدلالـة علـى اقسـام وفضـاءات الكليـة



المصطلحات	
1	حقول
2	بساتين
3	مبازل
4	بيوت لاستكية
5	ابنية تعليمية
6	دور سكنية
7	مدخل الحي السكني
8	مدخل الكلية الرئيسي
9	مدخل الطلبة
10	طريق ابو غريب
11	مضخه ري رئيسية
12	الاقسام الداخلية

(خارطة -١) موقع كلية الزراعة – جامعة بغداد

أما بالنسبة إلى مدخل الطلبة فهناك موقف خاص لسيارات الخطوط المخصصة لنقل الطلبة وهذا الموقف خارج الكلية فلا يسمح بالدخول لهم داخل موقع الكلية ، المدخل يفتقر الى شاخص للدلالة على نقطة الدخول ، كما انه غير محمي مناخياً من اشعة الشمس يمكن الطلبة من انتظار الخطوط الخاصة لنقلهم ، اما المدخل الخاص بالحي السكني فهو مغلق لاسباب امنية حالياً . مسارات الحركة المتمثلة بمجموعة من الطرق المتعامدة وهي رديئة من ناحية الاكساء ومعظمها لا يوجد على جوانبها ارصفة لسير المشاة كذلك عدم وجود فتحات خاصة بالمجاري على جوانبها لتصريف مياه الامطارمما يسبب تجمع المياه فيها كذلك فان وجود المناطق الترابية على جوانب الطرق يسبب في تحويلها عند سقوط الامطار الى منطقة طينية ، وكذلك عدم وجود الميلا ن المنتظم نحو الجوانب في بعض الطرق يسبب تجمع المياه فيها . توجد خمس مواقف مخصصة للسيارات الاول قرب مدخل الطلبة وهو اكبر المواقف ولكنه يمتاز بارضيه ترابية مما يسبب في تحويله الى منطقة طينية عند سقوط الامطار والثاني قرب العمادة والثالث امام القسم الرياضي والرابع قرب قاعة التاميم ، والخامس قرب النادي . وفيما يخص المساحات المكشوفة التي يتوجه اليها الطلاب وقت الفراغ وما بين المحاضرات او بعد انتهاء الدوام والتي تشمل : الحدائق ، الملاعب الرياضية، النادي والمطعم ،المسرح (قاعة التاميم) . ففي الكلية العديد من الحدائق يقع بعضها امام التسجيل ومجاور القاعات الدراسية ومقابل النادي وما بين الاقسام وامام العمادة وتشتمل على العناصر النباتية والعناصر الانشائية (تماثيل ، مساطب) والواقع الحالي لها هو الاهمال بشكل واضح اضافة لذلك فليس هناك تصميم مدروس لاي من حدائق الكلية بل تم زراعة الاشجار بشكل عشوائي ، فضلا عن ترك الاشجار والنباتات بدون قص او تشكيل او تقليم ورعاية مستمرة يسبب فقدانها لقيمتها الجمالية والوظيفية وكما موضح في (شكل - ٣) و(شكل -٤) . كما ان قله المماشي الرابطة بين اجزاء الحدائق وعدم وجود اماكن الجلوس المظللة الكافية يقلل من الاداء الوظيفي لها ، وتشمل الملاعب الرياضية ساحات لكرة القدم والسلة والطائرة والتنس وضم الملعب الرياضي الخاص بمعهد الادارة الى كلية الزراعة تسبب قي تشتيت الملاعب الرياضية وعدم تركيزها كما كان في منطقة محددة حيث يتعارض موقع هذا الملعب والفعاليات التي تجري فيه مع الفعاليات التعليمية القريبة منه . اما النادي والمطعم يقعان في بناية واحدة قرب مدخل الطلبة ويمتاز النادي والمطعم بزخم عالي من الطلبة وذلك بسبب استعماله كمكان يوفر خدمات الطعام والشراب وكمكان ترفيهي مهم ،كما تم انشاء كافتيريا خاصة بالاساتذة وطلبة الدراسات العليا ضمن بناية الدراسات العليا ، ويوجد مجموعة من الاكشاك بالقرب من النادي والمطعم ومبنى التسجيل وقسم الصناعات ، الاكشاك مصنوعة من الخشب وهي بحالة رديئة من ناحية بنائها وشكلها ولذلك فهي فاقدة لكل القيم الجمالية . اما مبنى المسرح (قاعة التاميم) يقع خلف التسجيل كان في السابق يستعمل لغرض احياء الحفلات واقامة الندوات وعرض الافلام السينمائية وفي الوقت الحالي مغلق بسبب حاجته الى اعمال الصيانة . اما المساحات المكشوفة ذات الوظيفة التعليمية والبحثية فتشمل حقول التجارب وبساتين النخيل والزيتون والمنحل والظلل الزجاجية والخشبية وحقول الدواجن والاغنام والابقار . وفيما يخص نظام الري في كلية الزراعة فهو السيح(الغمر) والابار الارتوازية ، حيث يستعمل لري الحدائق وما تشملة من مسطحات خضراء واشجار وشجيرات وحقول التجارب ، ولشحه الماء المتأتي من نهر ابو غريب الذي يمر محاذيا لحدود الكلية الشمالية واعطاءه باوقات غير منتظمة (الرشن) فقد تم حفر العديد من الابار الارتوازية لري الحدائق وحقول التجارب * . أن طريقة الري السيحي والابار هي طريقة قديمة وتسبب الكثير من المشاكل أهمها مشكلة الملوحة ، وذلك بسبب تبخر الماء وترسب الملح على سطح التربة وان مشكلة الملوحة تصبح اشد عندما تكون المياه الجوفية ملحية او عندما تصبح ملحية ، كذلك فان هذه الطريقة تسبب في عدم انتظام الري فعند بداية الري يكون قسم من مواقع التربة بحالة مشبعة بينما القسم الاخر الذي يكون عند نهاية دوره الري في حالة قريبة من نقطة الذبول ، وتشتد هذه المشكلة عندما تكون الارض غير مستوية مما يسبب تجمع الماء في المناطق المنخفضة وعدم وصوله الى المناطق المرتفعة ، فضلا عن ان الابار تحتاج الى مضخات تعمل بالكهرباء وانقطاع الكهرباء لساعات طويلة يسبب توقف هذه المضخات . ان مخطط الري يلعب دور مهم في مشاريع تصميم وتنسيق الحدائق ، وذلك من خلال وظيفته الجمالية والبيئية فضلا عن وظيفته في توفير مياه الري الضرورية لادامة النباتات .

* مقابلة شخصية لمسؤل الوحدة الهندسية في كلية الزراعة المهندس رياض خير الدين بتاريخ ٢٠٠٧/٨/٢٥

(شكل -٣) جانب من الواجهة الامامية لنادي الطلبة في كلية الزراعة

منظر لاحد الحدائق في كلية الزراعة ومقاعد الجلوس غير المظللة والمحاطة بالادغال

(شكل -٤)

(شكل -٥) جانب من الواجهة الامامية للمكتبة المركزية في كلية الزراعة والاهمال الواضح للمساحات المحيطة بها

(شكل -٦) منظر الى احد المبازل في كلية الزراعة وكثره الادغال يقلل من القيمة الوظيفية والجمالية

ان معظم السواقي الحالية وخاصة تلك التي توصل الماء من نهر ابو غريب الى أجزاء المواقع المختلفة غير مبطنة لذلك نلاحظ انتشار الادغال فيها وعلى اكتافها مما يعرقل حركة الماء ويقلل من كفاءتها الوظيفية فضلا عن فقدانها قيمتها الجمالية وكما موضح في (شكل -٥) ، توجد في بعض مساحات الكلية بعض السواقي المبطنة والتي تحولت هي الاخرى الى مكان لتجمع المياة والاوساخ ، وذلك بسبب عدم وجود حركة المياه فيها وتجمع مياه الامطار فيها من على الشوارع المحاذيه لها . ان مخطط الري الحالي يتطلب جهد اكبرمن ناحية الصيانة التي تشمل عمليات التنظيف من الادغال واعاده فتح بعضها وكذلك فان صعوبة التنظيم والادارة في تجهيز الماء من مصادره من نهر ابو غريب حالياً شبه معدومة وكذلك

بالنسبة للابار الارتوازية . ان المبازل الحالية مكشوفة وغير مبطنة مما ادى الى انتشار الادغال فيها وخاصة القصب الذي يعرقل حركة الماء ويقلل من قيمتها الوظيفية فضلا عن انه يسبب تدني القيمه الجمالية للموقع وكما موضح في (شكل -٦) . ان عدم انتظام عمل المضخات ادى الى ركود الماء في المبازل وتحويلها الى ما اشبه بمستنقع الماء الراكد وانتشار الكثير من الحشرات ومنها البعوض ، اضافة الى ذلك فان تسرب مياه المجاري من الحي السكني اليها سبب في انتشار الرائحة الكريهة من هذه المبازل . ان كفاءة هذه المبازل الوظيفية يقتصر على الاماكن القريبة منها وهذا ما أظهرته دراسات مناسيب المياه الجوفية في الموقع وهذا يعود الى غياب المبازل الحقلية .

٤- الجانب العملي :

تم توزيع استمارتي استبيان النموذج الاول خاص بالاساتذة والمختصين في المجال الزراعي والحيواني في الكلية وبنسبة ٨٠% من المتواجدين حاليا (خمسون استمارة) والنموذج الاخر خاص بالطلبة وبعدد مئة أستمارة لمعرفة اراءهم حول الموضوع ، وفيما يخص الاستبيان الخاص بالاساتذة والمختصين تضمنت الاستمارة مجموعة من الاسئلة على النحو الاتي :

السؤال الاول :

ما هي أهم السلبيات الاكاديمية والتطبيقية لموقع الحدائق والساحات المكشوفة في كلية الزراعة ؟

من خلال نتائج الاستبيان والتحليل الشخصي للباحث كان رأي المختصين المتفق عليه جميعاً هو :-

- ان المساحات المكشوفة بحاجة الى أعادة نظر في هيكلتها شكلاً ومضموناً فبعض المساحات تحتاج الى اعادة تصميم لاجل تحسين اداءها الوظيفي والجمالي والبيئي ، والبعض الاخر غير مستغل ويحتاج الى التأهيل لغرض الاستفادة منه .
- المسطحات الخضراء عبارة عن مزيج من الثيل وادغال لا حصر لها
- ان مياه الابار الارتوازية هو ماء نسبة الملوحة فيه E.C تتعدى 2.4 % .
- نظام الري المستخدم عن طريق غمر التربة بالماء طريقة قديمة ، وان تسوية الارض غير الجيدة تتسبب في تجمع ماء السقي في اماكن وعدم الوصول الى اماكن اخرى . وافضل طرق السقي وتلطيف الاجواء التي ينصح باستخدامها هو المرشاة المطرية التي تعطي النباتات ما يحتاجة من رطوبة فضلا عن غسلها للاوراق التي تجدد نظارتها .

كما ذكر المختصين ان أهم المسببات في تدهورالحدائق هو عدم اشراف جهه اكاديمية على الحدائق مثل قسم البستنة وعدم اشراك الاساتذة المختصين بخبراتهم وتجاربهم الاكاديمية في تطوير الحدائق ، واغلبها يحتاج الى اعادة برمجة وتخطيط ولابد من استحداث شعبة خاصة بالحدائق تقوم بالتنسيق مع العمادة لاعادة هيكلة هذه الحدائق وضمان التخصيصات المالية لذلك واعداد برنامج علمي فني في تصميم وترتيب وزراعة الحدائق لتحسين الاداء الوظيفي كل حسب موقعه واعداد برنامج علمي لمكافحة الادغال واستخدام اساليب حديثة ومتطوره لانظمة الري لادامة النباتات .

السؤال الثاني :

هل المساحات المفتوحة المخصصة حالياً لكلية الزراعة كافية أم انها بحاجة الى تغطية أضافية ....؟ ما هو اقتراحك للتطوير في هذا المجال ؟

لو تم احتساب المساحة الكلية للمساحات الخضراء حالياً في الكلية واجرينا حساب لحصة المنتسب الواحد من طلبة وكادر تدريسي واداري لوجدنا بان المساحة غير كافية ومنخفضة قياساً الى معايير المساحات الخضراء لحصة الفرد الواحد والتي قد تزيد عن ٢٨م ٢/شخص كمعيار معتمد لبعض الدول المشابهة مناخياً للعراق (التقرير الانمائي الشامل لمدينة بغداد حتى سنة ٢٠٠٠/ اب / ١٩٧٣/ ص٧٢) وانها بحاجة الى تغطية باضافة مساحة خضراء وعلى العكس مما يجري من قطع عشوائي لكثير من الاشجار والنباتات. ومن جملة ما قدم المختصون من مقترحات للتطوير :-

- وضع خطة عمل من قبل الاقسام المختصة وبشكل لجان او فرق عمل جادة ونشطة لاعادة تصميم الحدائق وتحسين ادائها .
- والاستفادة من الخبرات القديمة في مجال زراعة المحاصيل وتعويض الاشجار المصابة والمتساقطة بين حين واخرمع وضع جدول للجدوى الاقتصادية بذلك وعلى المدى الطويل .

- الحاجة الى التطوير بشكل علمي لخدمة الجانب العلمي والاغراض البحثية فضلا عن الجانب الجمالي ، وعلى سبيل المثال مداخل الكلية تحتاج الى تعريف اكثر كزراعة الزهور كالروز والجهنمي بالوانه المتعددة والاشجار الدائمة الخضرة لتعطي انطباع بانها كلية زراعة ، كما ان هناك ضرورة لاقامة المنشات من حدائق مائية ، بحيرة اسماك ، حديقة نباتية ، حديقة صخرية ، منشات زراعية مستغلة مثل البيوت الزجاجية والظلل الخشبية وغيرها وكل ذلك سيكون من اجل خدمة الجانب العلمي ويوفر بيئة علمية وبحثية يستفاد منها طلبة الدراسات العليا وكذلك الدراسات الاولية ولا يتم ذلك الا بعد توفر المستلزمات الاساسية لانجاح زراعة الغطاء النباتي من منظومة الري والتخصيصات المالية .
- ومن مقترحات التطوير زراعة المساحات المكشوفة امام قسم الوقاية والمحاصيل الحقلية والمساحات التي تمتد بين الظلل الخشبية قرب المسجد ولغاية الشارع العام ، زراعة الجزرات الوسطية بالورود على ان تنسق الشتلات بحسب تناسق الالوان ، كذلك زراعة الجزرات الوسطية لشارع قسم الصناعات الغذائية على امتداده .
- اختيار احد البيوت الزجاجية وكذلك الظلل الخشبية لزراعة اصناف محلية استوائية ، او من المناطق المعتدلة لتكوين حديقة نباتية Botanical garden فضلا عن استخدامها لاغراض التدريس وغيرها من التطبيقات العلمية والبحثية كما هو متبع في جامعات العالم .
- ادخال المكائن والالات الزراعية الحديثة لاغراض الحراثة والتعديل .
- اما مقترحات التطوير في المجال الحيواني فهناك الحاجة الى بناء حقول دواجن اضافية على نمط حديث فالموجود قديم جداً بالنسبة للتطور الحاصل في الدول العربية والعالمية . فضلا عن تطوير وتاهيل الحقل الحيواني وزراعة المساحات المتروكة بالعلف الاخضر لحل المشكلة التي يعاني منها الحقل الحيواني بصفة مستمرة لتوفير الغذاء للحيوانات الموجودة ضمن الحقل الحيواني والتي يستفاد منها لاجراء البحوث وتدريس الطلبة فضلا عن مردودها الاقتصادي في تزويد منتسبي الكلية بالحليب وكذلك تزويد قسم الصناعات بمادة الحليب لتصنيعها لانتاج الالبان .

السؤال الثالث :

كيف يتم الاستفاده من الحدائق والمساحات المفتوحة للأعمال البحثية حالياً وعلى مستوى الاساتذة الباحثين وطلبة الدراسات العليا ؟

ما هي مؤشرات التطوير التي يمكن ان تقترح من قبلكم ؟

حصيلة ما اوردته الاجابات الاستبيانية هو انه يمكن الاستفادة من الحدائق والمساحات المفتوحة من خلال استغلال بعضها لاغراض بحثية من قبل قسم البستنة لتطوير حدائق الكلية في مجال نباتات الزينة وعمل مشتل خاص لتربية نباتات الزينة بمختلف انواعها ، كذلك يمكن استثمار دروس مشاريع بحث التخرج لطلبة الصفوف المنتهية للاقسام العلمية ذات العلاقة وكذلك دروس الممارسة الحقلية للمرحلة الثالثة. وبالامكان تخصيص مساحات من الاراضي وزراعتها بمختلف المحاصيل بمشاركة الطلبة والاساتذة لكي يتدرب الطالب على كيفية اتمام عملية الزراعة وادارة المزرعة والمحصول وتشخيص الامراض وكيفية مقاومة هذه الافات والقضاء عليها . وبالامكان توظيف بعض بحوث طلبة الدراسات العليا في تطوير حدائق الكلية . وفي المجال الحيواني بالامكان استغلال القاعات الموجودة في حقل الدواجن لتربية الدجاج البياض وفروج اللحم خلال تجارب الاساتذة الباحثين وتجارب طلبة الدراسات العليا على ان لا تكون هناك محددات من حيث الربح المستحصل وان يكون للجامعة جزء من الربح الناتج من المشروع وكذلك الحال بالنسبة لحقول الحيوانات الكبيرة كالاغنام والابقار . يمكن توظيف بحوث طلبة الدراسات العليا في تطوير حدائق الكلية ودراسة المشاكل التي تعاني منها هذه المساحات فالبحث هو مشكلة وايجاد الحل لها ، كما يمكن الاستفادة من طاقة الطلبة في الدراسات الاولية من الاكثار من زراعة الشتلات لسد حاجة الحدائق والقيام بحملات زراعة بين حين واخر لتنمية الثقافة الزراعية في نفس الطالب وتحبيب العمل له وتنمية الشعور بان الكلية هي بيته الثاني ولابد من مساهمته في بنائها . يمكن عمل خارطة نباتية لمحتويات حدائق الكلية من اشجار وشجيرات ونباتات اخرى مؤشر عليها الاسماء العلمية واضافة نباتات اخرى غير متوفرة في حدائق الكلية فيصبح لدينا مختبر نباتي جيد يستفيد منه الطالب

والاستاذ في المجالات البحثية التي تخص النبات كدراسة ملائمتها وادائها الوظيفي ومسائل اخرى لها علاقة بالغابات والمردود الاقتصادي والنباتات الطبية وغيرها ، وبعد عمل دراسة مستفيضة عن تركيبة الحدائق ووضع المخططات الاستبيانية والارشادية عن محتوياتها يصبح لدى كلية الزراعة حديقة نباتية قادرة على ان تعطي مردودات علمية بحثية على المدى القصير والطويل .

السؤال الرابع :

هل يمكن تطوير المناطق الزراعية في الكلية وكذلك مناطق الانتاج الحيواني وتربية الاسماك وتربية النحل لتصبح مورداً مالياً مهما للكلية ..؟

ما هو الاسلوب الامثل الذي تقترحه للتوصل الى آلية عمل مثلى في الانتاج .

الاحصاء الاستبياني أكد أنه يمكن تطوير المناطق الزراعية في الكلية وكذلك مناطق الانتاج الحيواني لتصبح مورداً مالياً للكلية ، وللكلية تجربة ناجحة في سنة ١٩٩٣ وما بعدها حيث كان للمكتب الاستشاري الدور المميز في هذا المجال لان توفير الاموال هو الفيصل في ادامة وتطوير اي مشروع ، وتجربة المعمل في قسم الصناعات الغذائية خير مثال لما يعود من فائدة ومردود مالي للكلية فضلا عن تعليم وتدريب الطلبة . والاسلوب الامثل الذي اقترحه بعض المختصين :-

- ان تطوير حدائق الكلية من زراعة الازهار ينعكس على تطور اعمال المنحل لتوفر الازهار التي يحتاجها النحل كغذاء .
- يمكن استغلال البيوت البلاستيكية في انتاج الخضر ويمكن ان تكون انتاجية وتعليمية ،وفي قسمي الوقاية والبستنة يمكن استغلال البيوت البلاستيكية في انتاج بعض نباتات الزينة ويمكن ان تستغل للبيع وتطوير البيوت الزجاجية .
- زراعة الاراضي المتروكة بالخضر الموسمية وبيعها للمنتسبين والطلبة وتوفير مردود مالي للكلية ويمكن اشراك الطلبة والمنتسبين باسهم تعاد لهم بارباح سنوية .
- انشاء بساتين متكاملة متخصصة مثل النخيل والزيتون والعنب والحمضيات وغيرها .
- تربية الدواجن وانتاج بيض المائدة .
- تربية الاسماك بشقيها الزينة والانتاجي .
- توسيع معمل الالبان ليشمل منتجات مختلفة .

اما حقول الانتاج الحيواني فهي بحاجة الى تطوير واستبعاد الحيوانات المصابة ورش الحقول وتعقيمها من الاصابات لكونها تعرضت لاصابات عديدة . اما حقول الدواجن فالقاعات بحاجة الى تطوير واصلاح خاصة القاعة الارضية للبياض واصلاح السقوف الثانوية فيها كي تصبح صالحة للتربية ولابد من تشكيل فرق عمل من ذوي الخبرة لغرض اكمال احتياجات الحقل من النواقص ووضع خطة لاتمام عملية التطوير من خلال دراسة واقع المشاريع السابقة في الكلية وايجاد دعم لصندوق التعليم او اي تنظيم مالي اخر بحيث يمكن الاستفادة منه في تاسيس المشاريع الانتاجية ذات الطبيعة الانتاجية المكثفة .

الاسلوب الامثل هو وضع لجنة في كل قسم بحيث تكون مسؤولة عن اي مشروع يتم عمله وتحدد كل الاسس الرئيسية للمشروع سواء في الدواجن او الاغنام او الاسماك وتكون على دراية بعدد الحيوانات وكمية العلف ومتابعه المشروع بصفة مستمرة وتكون على دراية حتى بالتسويق والربح المتحصل على ان يقسم الربح بنسب معينة وجزء منه يكون للكلية ولا تاخذ الجامعة الارباح بحيث تكون حصة الارباح لمنتسبي الجامعة اكثر من حصص منسبي الكلية .

ومن المهم التخلص من الروتين الاداري والمالي وربط المشاريع بالاقسام العلمية ومن خلال لجان يمثل فيها الادارة والحسابات . ويمكن اشراك الطلبة والموظفين بأسهم يمكن شرائها سنوياً ويخصص لها من ارباح هذه المنشات سواء كانت زراعية او ثروة حيوانية ويكون هذا عرف سنوي يضاف لموارد الكلية الموجودة.

السؤال الخامس :

أي افتراضات تجدونها مهمة في هذا المجال ؟

مثلا – التوسع في استخدام المسطحات المائية - انشاء مراكز بحثية – اقامة معارض للزهور والنباتات – انشاء حدائق نباتية مكشوفة ومغلقة .

كانت اجابات المختصين حول السؤال الاخير ان جميع الافتراضات مهمة ولكن هناك معوقات لتنفيذ كل منها ، فهناك عوامل رئيسية لانجاح اي مشروع سواء كان لغرض ترفيهي او انتاجي او تعليمي وهذه العوامل تتلخص في :

- التخصيصات المالية الكافية ، مصادر المياه (شبكات الارواء) ، الايدي العاملة – ذات الخبرة والممارسة العلمية الكافية ،ان يتولى امر هذه المشاريع من له دراية وفهم لاهمية هذه المشاريع وهذه العوامل غير متوفرة في الكلية .

- فالمسطحات المائية مهمة جداً بالنسبة للحدائق المفتوحة ولكنها تتطلب ادارة بيئية جيدة لكي لا تصبح مستقبلاً بؤرة للمياة الاسنة والامراض ، وهناك حاجة ماسة لها حيث ان طلبة الدراسات الاولية يعانون من عدم وجود مسطحات مائية ولو بحيرات صغيرة يمكن استغلالها في ابحاث الطلبة في قسم الثروة الحيوانية كبحيرة اسماك او دراسة الاحياء المائية الاخرى .

ان أنشاء المراكز البحثية هو مطلب جيد ولكنه يحتاج الى أدارة مهنية وقد تم انشاء مراكز بحثية في الكلية مثل مركز المكافحة الاحيائية ومركز بحوث النخيل وتمت الموافقة على انشاء مركز بحوث النباتات الطبية ولكن المشكلة في التخصيصات المالية وعدم وجود الملاك . علماً بأنه من الممكن الوصول الى كل هذه الامكانات المذكورة من خلال التغطية الذاتية للمشاريع كامكانات تجارية قابله للتسويق .

كما بين بعض المختصين في الكلية أهم معوقات تطوير وأعادة تاهيل المساحات المكشوفة والخضراء ويمكن تلخيصها في :

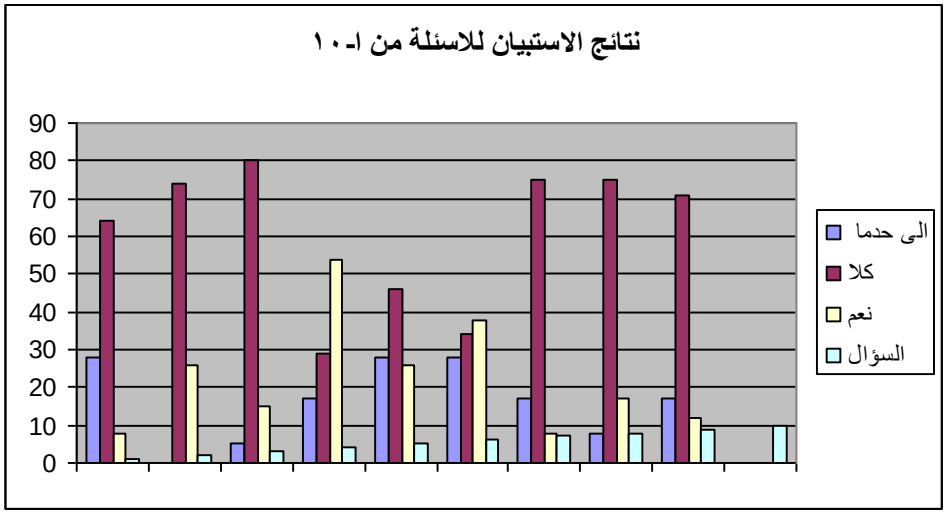
- عدم توفر ماء الري الضروري لادامة المسطحات الخضراء والذي يمكن توفيره من خلال تطوير نهر ابو غريب .
- عدم توفر التخصيصات المالية لشراء المستلزمات الزراعية والتي ممكن توفيرها من خلال المردود الاقتصادي الكبير للمشاريع المقترحة .
- عدم استخدام المعلومات الاكاديمية في جوانب تطبيقية .
- عدم استخدام المكننة وخاصة في مجال الادغال وعدم استخدام المبيدات الكيمياوية ووسائل المكافحة الاخرى .
- عدم وجود اشراف لمختصين على الحدائق وقلة الايدي العاملة والتي ممكن تأمينها من خلال عائدات الارباح .
- الانقطاع المستمر للتيار الكهربائي يؤثر على عمل مضخات الري والسقي مما قد يترتب عليه انشأ محطة كهربائية خاصة .
- عدم وجود المتخصصين في هندسة الحدائق والاشراف على ادامتها ومن الممكن جلب اختصاصين من خارج القطر .
- عدم استخدام التقنيات الحديثة في هندسة الحدائق والذي قد يترتب عليه صرف رأس مالي اولي قابل للتسديد بسهوله مستقبلاً ومن خلال ضمانات مصرفية .
- عطل مضخات الري والاعتماد على عمال ذوي كفاءة محدودة في التشغيل وما اسهل تغطية هذه المشكلة والذي يتطلب المشاركة الجماعية لجميع المختصين في هذا المجال .
- عدم وجود برنامج علمي فني في تصميم وترتيب وزراعة الحدائق والساحات .
- عدم استغلال المساحات الكبيرة في حدائق الكلية وحقول الدواجن والاغنام بالشكل الامثل حيث تحتاج هذه المساحات الى اعادة برمجة وتخطيط على سبيل المثال مداخل الكلية يفترض تظهر من خلال زراعة الزهور كالروز والجهنمي بالوانه المتعددة والاشجار والخضرة لتعطي انطباع بانها كلية زراعة .

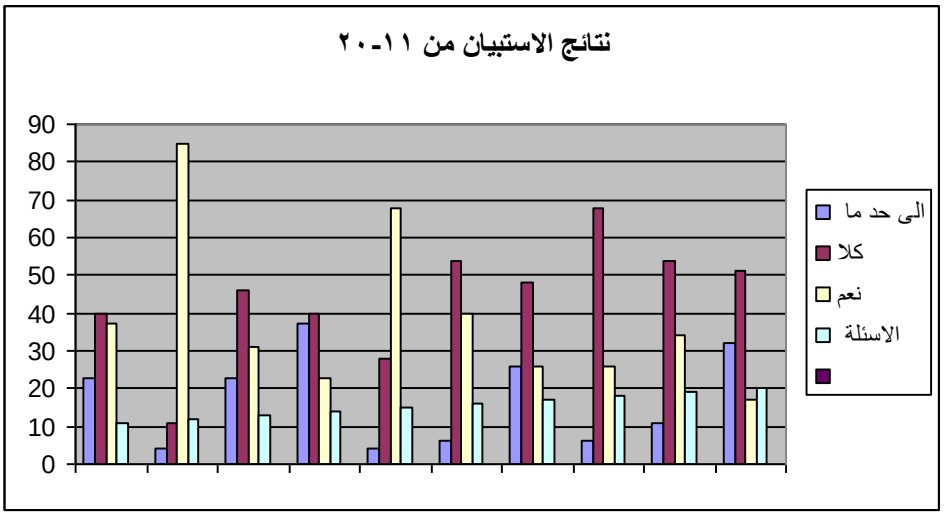
النموذج الثاني لاستمارة الاستبيان : خاص بالطلبة حيث تم توزيع عدد منها على الطلبة تضمنت الاستمارة مجموعة من الاسئلة حول تخطيط وتصميم الفضاءات الخارجية والحدائق لكلية الزراعة ومعرفة اراءهم حول اهم السلبيات

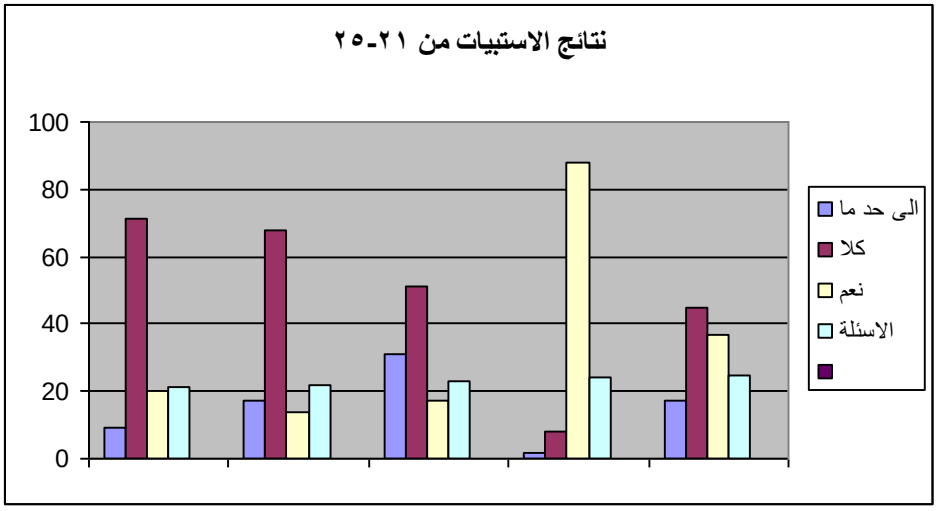
(جدول -١) محصلة نتائج استمارة الاستبيان الخاصة بطلبة كلية الزراعة مع نتائج الاستبيان

	أسئلة حول تخطيط وتصميم الفضاءات الخارجية لكلية الزراعة	نعم	لا	الى حدما	نتائج الاستبيان
١	مدخل الكلية يمنح شعور بالترحيب بالقادمين والوضوح والحماية والسيطرة .		*		المدخل لايمنح شعور بالترحيب والوضوح والحماية والسيطرة
٢	توفر مواقف سيارات مسقفة لحماية السيارات من الاكتساب الحراري .		*		مواقف السيارات غير محمية من اشعة الشمس .
٣	ممرات الحركة التي تربط كافة الاقسام مسقفة مكونة مناطق مظللة للتحرك والتجمع الطلابي		*		ممرات الحركة ومناطق التجمع غير مسقفة .
٤	توفر العلامات الدالة للارشاد الى اقسام الكلية			*	تتوفر العلامات الدالة الى الاقسام .
٥	توفر فضاءات للتحرك والتجمع والراحة والقراءة مع توفر كافة الخدمات الخاصة بها		*		لا تمتلك الكلية فضاءات للتجمع تتوفر فيها كافة الخدمات
٦	توفر فضاءات لممارسة فعاليات معينة كالطعام والشراب			*	تتوفر فعاليات معينة كالطعام والشراب
٧	فضاءات التجمع غنية بالمعالجات التصميمية		*		لا توجد فضاءات غنية بالمعالجات التصميمية
٨	وجود علامات وارشادات دالة على مواقف السيارات		*		لا توجد علامات وارشادات دالة
٩	استخدام المبلطات ذات الاشكال الجذابة		*		لم يتم استخدام المبلطات ذات الاشكال الجميلة
١٠	تشكل الملاعب عنصر مهم في الفضاءات الخارجية اذ انها ضرورية لاحتفاظ الطلبة بنشاطهم ولياقتهم البدنية			*	نعم الملاعب تشكل عنصر مهم في الفضاءات الخارجية
١١	توفر فضاء خاص لاقامة المناسبات الوطنية او مناسبات التخرج يستوعب اعداد كبيرة من المحتفلين		*		لا يتوفر فضاء خاص لاقامة المناسبات الوطنية او مناسبات التخرج يستوعب اعداد كبيرة من المحتفلين
١٢	هل ترغب بوجود تصميم مدرج ومنصة للخطابات مع نظام انارة ونظام مكبرات صوت واماكن لوضع الاعلام واللافتات .	*			نعم يرغب جميع الطلبة بوجود تصميم مدرج ومنصة للخطابات
١٣	توفر نقاط جذب مثل نافورة ، اماكن للجلوس ، استعمال تدرج بالمستويات ضمن فضاءات تجمع الطلبة		*		لا تتوفر نقاط جذب
١٤	توفر فضاءات هادئة للراحة والتأمل والقراءة بعيدة عن شرايين الحركة مزودة بمقاعد جلوس مظللة		*		لا تتوفر فضاءات هادئة للراحة
١٥	وجود ساحة لكرة الطائرة وللسلة وللتنس	*			نعم توجد ساحة لكرة الطائرة وللسلة وللتنس
١٦	وجود ساحة مركزية لكرة القدم والالعاب السويدية		*		لاتوجد ساحة مركزية لكرة القدم والالعاب السويدية
١٧	توفر اكشاك بيع المرطبات والصحف وغيرها		*		لاتتوفر اكشاك بيع المرطبات والصحف وغيرها
١٨	استخدام عنصر الماء لاضافة متعة بصرية وصوتية		*		لم يتم استخدام عنصر الماء
١٩	استخدام مقاعد للجلوس مريحة وجذابة		*		لم يتم استخدام مقاعد للجلوس مريحة وجذابة
٢٠	توفر اماكن تقوي الاحساس بالجانب الجمالي		*		لا تتوفر اماكن تقوي الاحساس بالجانب الجمالي
٢١	وجود المسقفات ذات الاشكال والالوان الممتعة التي توفر للطالب الراحة النفسية من عناء الدراسة في اوقات الراحة		*		لاتوجد المسقفات ذات الاشكال والالوان الممتعة التي توفر للطالب الراحة النفسية من عناء الدراسة في اوقات الراحة
٢٢	وجود تصاميم وتشكيلات مميزة من احواض النباتات والزهور الفريدة النوعية .		*		لا نوجد تصاميم مميزة من احواض النباتات والزهور الفريدة النوعية
٢٣	هل تستشعر بوجود تنوع في الفضاء عند تجولك في حدائق وساحات الكلية		*		لا يوجد تنوع في الفضاء
٢٤	هل ترغب بأن يكون لك دور ومساهمة في صيانة حدائق الكلية .	*			نعم يرغب الطالب بأن يكون له دور ومساهمة في صيانة حدائق الكلية .
٢٥	اللون هو أحد مظاهر الاستيعاب البصري هل تستشعر تلوين الفضاءات بقيم لونية تعطي ردود فعل فكرية وعاطفية (لون الزهور ، لون الارضية ، لون حاويات الزهور ، لون مقاعد الجلوس)		*		ليس هناك تلوين للفضاءات بقيم لونية تعطي ردود فعل فكرية وعاطفية (لون الزهور ، لون الارضية ، لون حاويات الزهور ، لون مقاعد الجلوس)

(جدول -٢) يوضح نتائج الاستبيان للاسئلة المشار اليها اعلاه .







- مناقشة وتحليل النتائج :

من خلال الاستبيان اتضح :-

١- ان الحدائق والمساحات المكشوفة في الكلية بحاجة الى اعادة نظر في التخطيط والتصميم لتحسين الاداء الوظيفي والبيئي والجمالي ، فالحدائق القريبة من القاعات الدراسية مهملة وغيرمستغلة وهي بحاجة الى تخطيط وتاهيل واستغلال لقربها من القاعات الدراسية ، فالطالب بحاجة الى اماكن تقوي الاحساس بالجانب الجمالي والصفاء الذهني والراحة النفسية ، ان عدم وضع التصاميم المدروسة للحدائق وقلة العناية بها سبب في تدهورها وفقدانها لعناصر المتعة والاثارة والجمال والاداء الوظيفي الجيد ، ويفترض ان تزود هذه المساحات باماكن الجلوس والمسقفات ذات الاشكال الجميلة والالوان الممتعة التي توفر للطالب الراحة النفسية من عناء الدراسة في اوقات الراحة .

٢- وفيما يخص ممرات الحركة التي تربط كافة الاقسام لوحظ من خلال الدراسة بانها تفتقرالى ابسط المقومات الوظيفية والبيئية والجمالية ، فاغلب الارصفة غير مبلطة وبحاجة الى مسقفات لتوفر مناطق مظللة للتحرك والتجمع الطلابي للحماية من اشعة الشمس صيفاً والمطر شتاءا .

٣- ً وكذلك اماكن وقوف السيارات بحاجة الى مسقفات لحماية السيارات من الاكتساب الحراري فضلاً عن عدم كفايتها فلا يوجد قرب المكتبة وبناية الدراسات العليا موقف للسيارات وهذا يسبب بعض الارباك ووقوف السيارات بشكل عشوائي وخاصة في اوقات معينة كان تكون هناك مناقشة لاحد طلبة الدراسات العليا .

٤- ومن خلال الاستبيان والملاحظة اتضح انخفاض القيمة الوظيفية والجمالية للنباتات من اشجار وشجيرات وزهور بسبب عدم العناية بها بصورة كافية فهي بحاجة الى التقليم والتشكيل والقص وتوفير ماء السقي بشكل منتظم والتربة الجيدة ، فهناك اهمال واضح وقص عشوائي للاشجار المتقدمة في العمر ، ويفترض ان تزرع شتلات سنوياً لتعويض التالف منها والمصاب .

٥- ان نظام الارواء الحالي في كلية الزراعة هو السيح (الغمر) والابار الارتوازية وملوحة ماء الري المستخدم في السقي (ماء نهر ابو غريب) ساهم بشكل واضح في رفع نسبة الملوحة في التربة ، فضلا عن ذلك كمية الماء غير كافية لسقي الحدائق لشحة الماء المتأتي من نهر ابو غريب واعطاءه بأوقات غير منتظمة (الرشن) ، كما ان اهمال السواقي وتركها بدون تنظيف بشكل مستمر من الادغال حولها الى مكان لتجمع الادغال والاوساخ ومياه الامطار ، ومن خلال المشاهدة تبين ان المبازل الحالية هي من المبازل المفتوحة غير المبطنة مما يسبب انتشار الادغال فيها كما ان انقطاع التيار الكهربائي لساعات طويلة وعدم عمل مضخات السحب بصورة منتظمة سبب في تحولها الى مستنقع من الماء الراكد وبالتالي فقدانها لقيمتها الجمالية والوظيفية .

٦- وفيما يخص البيوت البلاستيكية والمظلات فعدم وجود الدعم المالي والتخطيط الصحيح لاستثمارها وتطويرها ادى الى تلف ما موجود منها وبدلاً من التوسع والتطوير نجد الفشل والخسائر وكذلك المنحل ، وحقول الدواجن وحقول الاغنام وغابات الزينون والنخيل ، فكلية الزراعة فيها من الموارد والخبرات التي لو تم استغلالها وتوظيفها بالشكل الصحيح لكان هناك من المردود الاقتصادي الذي يعود على المنسبين والمنطقة المحيطة بالفائدة فضلا عن المردود العلمي الذي يعود على الطلبة والباحثين من طلبة الدراسات العليا والدراسات الاولية بالفائدة العلمية .

٧- وفيما يخص استمارة الاستبيان الخاصة بالطلبة فقد أشارت نسبة كبيرة من الطلبة فيما يخص مدخل الكلية بانه لا يمنح شعور بالترحيب للقادمين والوضوح والحماية والسيطرة . ولايعطي الاحساس للزائر من اول وهلة بانه داخل في كلية زراعة وذلك لفقدان العناصر التعريفية والتجميلية من احواض زهور مميزة ، ونباتات نادرة ، ولوحة كبيرة الحجم تتصدر الواجهة يمكن ذكر بعض المعلومات عن تاريخ تأسيس الكلية والاقسام التي تتضمنها .كما أشار معظم الطلبة بعدم توفر مواقف سيارات وممرات حركة مسقفة ، مما يؤثر على كفاءة استخدامها ، ان حماية الفضاءات من اشعة الشمس واتخاذ الحلول والمعالجات بهدف توفير بيئة مناخية ملائمة

تزيد من كفاءة الاستخدام ضرورة من الضرورات التي لا يمكن الاستهانه بها . ومن خلال الاستبيان تبين افتقار الحدائق وممرات الحركة الى كثير من المعالجات التصميمية والتخطيطية ولايوجد تنوع بالفضاء من ناحية الاستخدام والوظيفة . أشار معظم الطلبة أن أغلب الفضاءات والحدائق القريبة من اقسامهم لم يتم تخطيطها بحيث يتم الانتفاع منها من ناحية توفير الخصوصية اللازمة لمختلف الاستعمالات الخاصة كأن يكون جانب منها مخصص للاستراحة ، وجانب مصمم كمنطقة تجمع ، والقسم الاخر مصمم لممارسة فعالية الاكل والشرب ، والاخر مصمم للقراءة والتأمل، والاخر لفعالية الاستنساخ وشراء بعض اللوازم التي يحتاجها الطالب . ولغرض تحقيق ديمومة هذه الفضاءات فلا بد من وجود ادارة تتحمل مسوؤلية المتابعة والاشراف والصيانة ومحاسبة الطالب في حالة الاخلال بالقانون من ناحية المحافظة على النظافة او الاضراربالممتلكات. كما تبين حاجة الحدائق الى اماكن جلوس مريحة وجذابة مظللة ، ونقاط جذب كالنافورات والتماثيل والتشكيلات المميزة من احواض الزهور التي تفترض ان تزين ممرات الحركة ومداخل الاقسام لتعطي شعور واضح بخصوصية الموقع . كما اوضحت الدراسة بحاجة الكلية الى فضاء خاص لاقامة المناسبات الوطنية او مناسبات التخرج وأيد أغلب الطلبة فكرة وجود مدرج ومنصة للخطابات ، وان يكون للطالب دور ومساهمة في صيانة حدائق الكلية .

- الاستنتاجات الختامية :

- غياب التخطيط الصحيح والاستغلال الامثل للحدائق والمساحات المفتوحة .
- قلة المتخصصين في مجال تصميم وتنسيق الحدائق ، وعدم وجود شعبة فنية تخصصية تتولى عملية الاشراف والمتابعة والتخطيط ضمن برنامج علمي .
- حاجة الكلية الى التخصيصات المالية الكافية لتغطية نفقات التطوير والانشاء والادامة من مستلزمات زراعية ومكائن واسمدة وكافة المتطلبات .
- نظام الري المتبع قديم ولا يفي بحاجة الحدائق والبساتين والحقول لغرض ادامة النباتات .
- حاجة حقول الدواجن والاغنام الى التطوير والادامة لكي تتماشى مع التطور الحاصل في بلدان العالم .

٧- التوصيات الختامية :-

- وضع برنامج علمي فني لاعادة هيكلة المساحات المفتوحة والخضراء والاستغلال الامثل حيث تحتاج هذه المساحات الى اعادة برمجة وتخطيط ، من خلال جهة تخصصية تكون مسوؤلة عن ادارة وصيانة الحدائق .
- فتح مجال الاستثمار امام الاساتذة وذلك من خلال تخصيص بعض المساحات لزراعتها بمشاركة الطلبة لكي تكون مجال لتدريب الطالب على كيفية اتمام عملية الزراعة وادارة المزرعة والمحصول وتشخيص امراضة وحشراته وفي نفس الوقت تعود بالمردود المالي على الاستاذ ويمكن وضع نسبة قليلة من الارباح للكلية ضمن ا لية يتفق عليها .
- ايجاد نظام يحفز العاملين على العمل يتناسب مع جهودهم .
- اشراك القطاع الخاص بالاستثمار لتامين التخصيصات المالية .
- تامين مياه الري وذلك من خلال التنسيق مع دائرة ري ابو غريب لضمان ايصال الكميات اللازمة من الماء وزيادة عدد الابار المحفورة وتوفير المضخات لايصال كميات الماء التي تحتاجها الحقول .
- استخدام انظمة ري متطورة والاستفادة من تجارب الدول المشابهة لظروف العراق البيئية كما جاء في الدراسة .
- اقامة المعارض لبيع الزهور ونباتات الزينة بهدف تثقيف الزوار وتوعيتهم بيئياً وتنمية الذوق الجمالي والحسي .

- الاهتمام باعمال الصيانة والقص والتقليم والتشكيل لرفع القيمة الجمالية للحدائق .
- تشكيل لجنة للاشراف والمتابعة والتفتيش على سير العمل وتقديم تقارير دورية وحسب الحاجة .
- اعطاء الحقل الحيواني الى شركات مستثمرة والتعاقد معها لعدد من السنوات وفق الية معينة يتم الاتفاق عليها

المصادر

١. 1. www.momra.gov.sa/specs/guid0014.asp1 (دليل زراعة النباتات الملائمة لمشاريع التشجير في مناطق البيئة المختلفة) .

٢. www.momra.gov.sa/Specs/guid0021.asp (وزارة الشئون البلدية والقروية ، المملكة العربية السعودية ،اسس تصميم وتنفيذ وصيانة الحدائق العامة ، ٢٠٠٤)

٣. http://civilawy.jeeran,com/lectures2.htm (تنسيق الحدائق ، قسم عمارة ، كلية الفنون الجميلة ، ٢٠٠٤)

٤. www.uaeagri.ae/home.aspx?ctname=ArticleDetails.ascx&rid=76 (مقالات وبحوث ،هندسة الري الحقلي.)

٥. www.uaeagri.ae/home.aspx?ctname=NewsDetails.ascx&news_id=319 (القلاف ، عبد المحسن ، مدير الهيئة العامة لشؤون الزراعة ، جريدة الاي العام الكويت ، ٢٠٠٧)

٦. www.zira3a.net/forum/showthread.php?t=708 (البطراوي ، احمد ، قسم الاراضي والمياه ، شبكات الري بالتنقيط ، ٢٠٠٥.)

٧. www.fao.org/docrep/007/ad820a/ad820a03.htm/ (سياسات التنمية الزراعية في الشرق الادنى ، القضايا – المتطلبات – المنهجيات ، ٢٠٠٣) .

٨. www.urar.org.sa/ibda/mahawer3-4.htl (قيصران ، عبد بن محمد بن صالح ، البعد البيئي في تخطيط وتصميم مدينة ينبع الصناعية ، ٢٠٠٤)

٩. www.oqla.gov.sa/alashratat.htm. (الفراج، عز الدين ، تطوير خدمات مشاريع التشجير ، ١٩٨٦)

١٠. www.arriyadh.com/Tourism/LeftBar/Gardens

١١. مجلة المزارع الحديث ، ١٩٥٢

COMPUTER VIROLOGY: SELF-DEFENSE SYSTEM

Hamid M. A. Abdul-Hussain & Hamed Mizher Shabib
Computer Engineering Dept.
College of Engineering, University of Baghdad

ABSTRACT

Based on the biological human models in defending human body against viruses, a new approach in designing the anti-virus system is introduced. This approach is called SDS(Self-Defence System). The principle of the SDS is that each executable program is responsible of defending itself against viral-attacks. In this system, each executable program is injected with basic anti-virus component which is called Self-Defence Routine. This routine, together with dedicated anti-virus loading program are used to construct the SDS which protects the computer system from virus invasion.

INTRODUCTION

Today, the computer virus pandemic becomes a serious security threat to causal home computers and large corporate networks. Over the years, the anti-virus industry has had to keep pace, as virus writers have become more sophisticated. Therefore, the effort to combat the computer virus must continue until an ideal, universal anti-virus system is designed. Researchers have taken many approaches, and some of the newest and most promising anti-virus technology is modeled on the way the human body fights viruses [1].

Based on the similarities between human and computer viruses (both types of viruses latch onto a host, use its resources to reproduce, and cause a range of symptoms). The objective of this work is to build computer immune system. The proposed system is called Self-Defence System (SDS). The SDS mimics the characteristics of human immune system by distributing the anti-virus components through the computer in the same way the human immune system distributes the anti-bodies. The concept of the SDS is to vaccinate each executable program with special anti-virus component that will detect and eradicate any foreign code attached to the executable. The vaccinated executable is loaded and executed by special centralized anti-virus program which is designed to prevent virus infection and damage to the computer system.

In part this research has been a follow-up on the two papers; the first titled “Computer Virology: Formal Analysis of Computer Viruses” [7], and the second titled “Computer Virology: Toward Designing an Ideal Anti-virus System” [4]. Some terms, classifications, and concepts presented in this research are thoroughly explained in the above papers. For example the efficiency of the SDS is determined according to the analysis criteria described in the second paper mentioned above.

SELF-DEFENCE CONCEPT

The self-defense concept is that: **“Every executable program must be capable of protecting itself against viral attacks by detecting and eradicating any virus attached to its body”.** Accordingly, special part of each executable must be reserved for the self-defence task. This part is called **“Self-Defence Routine (SDR)” and** defined as: **“The part of the**

executable that is specifically designed to protect the executable against viral attacks". SDR is the first basic anti-virus component in SDS, which is embedded, as integral part of an executable E. E will be called the **"protected executable"** along this research. The second basic anti-virus component in SDS is a special OS program called "Load and Execute Program (LEP)". LEP is used to load executables into memory and execute their SDRs secretly by using a special secret communication protocols. LEP and SDR cooperate to satisfy the requirements of the IAVS.

The communication between LEP and SDR consist of two phases (PHASE-ONE and PHASE-TWO). During PHASE-ONE, LEP must transfer control to SDR without giving the attached virus "if any" the opportunity to execute. If SDR detect a virus infection, it will try to eradicate the attached virus and repair the protected executable using the Foreign Block Eradication Algorithm (FBEA). If FBEA failed, SDR must alert LEP to switch to PHASE-TWO. In PHASE-TWO, LEP will execute the virus in a virtual computer system. After the virus execution is completed, SDR will try to eradicate the virus and repair the infected executable using Virus Follower Eradication Algorithm (VFEA). The following sections will discuss how SDR and LEP are designed, and the communication protocols they use in PHASE-ONE and PHASE-TWO.

SECRET ENTRY POINT

Most viruses tacks themselves into an executable program and ensure that they will execute before their host, therefore, they must redirect the executable standard entry point to point to the virus entry point. The standard entry point is the location to which the OS transfers control when deciding to execute the executable. The standard entry point is considered the most vulnerable spot for viral attacks.

In PHASE-ONE, LEP must execute SDR without giving the opportunity to any attached virus "if any" to execute. Clearly, this cannot be done by transferring control to the standard entry point of the executable because viruses know where the standard entry point of the executable is, and they always redirect it to ensure that they will get the opportunity to execute before the infected executable. There is only one way to ensure that SDR will get the opportunity to execute before any attached virus. That is, by using a nonstandard or secret entry point. This kind of entry point is called trapdoor: "**The Trapdoor is a secret entry point to the executable [2]".** Fig.1 shows how the trapdoor can be used.

In fact, the secret entry point will be used to solve the standard entry point problem. Therefore, the trapdoor location must vary from one executable to the other. For example, while the trapdoor of the executable 'E1' is located at offset 'X1', the trapdoor of the executable 'E2' must be located at offset 'X2', where 'X1≠X2'. Because the same standard entry point problem will be faced again, if the trapdoor location is standardized for all executables. For example, if the trapdoor location is standardized to be at offset 'X1' for all executables, then: **"The virus will know that there is another entry point to the executable located at offset 'X1'. Consequently, it will redirect to this entry point to ensure that it will get the opportunity to execute before its host".** Therefore, a new technique that will enable LEP to find or mark the trapdoor while, at the same time, viruses cannot find the trapdoor. The following sections describe how LEP mark the trapdoor, the standard SDR module format, and how different LEPs can exchange protected executables between each other.

TRAPDOOR MARKING

LEP can find the trapdoor location of a given executable by using special type of information stored in the executable itself. This special type of information will be called the "**Trapdoor Mark"** and defined as "**A special code or string of characters whose location (LOC) or value (TMV "Trapdoor Mark Value") can be used to find the trapdoor".**

LEP must calculate the TMV and store it into LOC using criteria known by LEP only. This criterion differs from one LEP to the other. Therefore, viruses cannot estimate in advance which trapdoor mark is used by a randomly selected LEP. For example, if LEP1 and LEP2 are used in two different computers PC1 and PC2 respectively. LEP1 define (TMV= "HMS"), and

define the first byte position which follow the 3'rd character as the trapdoor (Trapdoor=LOC+3), while LEP2 define (TMV=666) and (Trapdoor=LOC+2). Since the virus cannot estimate in advance in which computer it exist, in PC1, PC2, or some other computer, it cannot estimate in advance whether the TMV is, "HMS", 666, or any other number or string. The efficiency and reliability of the secret entry point depend on the TMV type and the method used to generate it. In general, two TMV types can be defined:

FIXED TMV

In this type, the same TMV used for all executables in the system. The generation and searching formulas are very simple in this case. There is, however, a vulnerable spot in this case. Since the TMV is used in all executables, it will be the common code between all executable. Viruses can determine the fixed TMV by comparing two or more executables to see which code is common between them. This code might be the TMV. Note that if each executable in the system uses the suggested self-defence technique efficiently, no virus will get the opportunity to execute and therefore it cannot perform the comparison described above.

Variable TMV

In this type, a different TMV used for each executable in the system. Only LEP knows how to generate and find this TMV for a given executable. Two general techniques can be used to generate a variable TMV:

A- Value-Based Technique: "In this technique, TMV is calculated using a predefined formula and stored in LOC". The general Value-Based generation formula can be defined as follow: **"TMV= f(g(EBV), LBC)".** Where:

- **EBV "Executable-Based Variable":** This value is constant with respect to a given executable. But it differ from one executable to the other. Therefore, LEP can ensure that TMV will change from one executable to the other.
- **LBC "Loader-Based Constant":** This value is constant with respect to a given LEP. But differ from one LEP to the other. This ensures that the TMV will change from one LEP to the other even if both use the same EBV, g, and f.
- **f and g :** Functions (arithmetic or logical relation).

In general, LEP must search for the trapdoor mark location (LOC) using TMV. The general Value-Based searching formula can be defined as follow: **"IF ([P]=TMV) THEN (LOC=P)".** Where:

- **P= Address Pointer**
- **[P]= Content of the location at address P.**

For example, the executable file name can be used as EBV, because it is constant for a given executable file. The TMV generation formula will be: **"TMV= f(g(Executable Name),LBC)".** Let us define the following, for LEP1 and LEP2:

Component	LEP1-Definition	LEP2-Definition
g	ASCII of the 1st Character + ASCII of the 2nd Character + ASCII of the 3rd Character	ASCII of the 2nd character - ASCII of the 4th character
LBC	6	305
f	g + LBC	g ⊕ LBC

Tab.1 shows the TMV generated for three different file names by using the definitions shown above. LEP1 can find LOC for HAMED.COM using the following algorithm:

Calculate :TMV=72+65+77+6=220
P=0
IF ([P]=220) THEN (LOC=P) ; GO TO 6
P=P+1
IF (P<Size of HAMED.COM) THEN 3

End.

B- Condition-Based Technique: "In this technique, TMV is selected so that a predefined condition satisfied". The general Condition-Based generation formula can be defined as follow: **"Select TMV To Satisfy: f(LOC,LBC1)=g(LOC,LBC2)".**

The general Condition-Based searching formula can be defined as follow: **"IF (f(P,LBC1)=g(P,LBC2)) THEN (LOC=P)".** For example, the following can be defined:

LBC1	f(P,LBC1)	LBC2	g(P,LBC2)	Condition
0	[P]+LBC1	0	P+LBC2	[P]=P
1	[P]+LBC1	2	P*LBC2	[P]+1=P*2
19	[P]-LBC1	0	P⊕[P-1]+LBC2	[P]-19=P⊕[P-1]

Tab.2 shows the TMVs generated to satisfy the condition "[P]=P+3" for the three executables shown in Tab.1. The following algorithm can be used by LEP to search for LOC, assuming that f, g, and LBC are defined as shown above:

P=0
IF ([P]=P+3) THEN (LOC=P) ; GO TO 5
P=P+1
IF (P<Size of HAMED.COM) THEN 2
End.

However, it is very important to test the influence of virus infection on the selected condition. For example, if the condition [LOC]=LOC+3 used, and a virus inserts 256-bytes in the start of HAMED.COM and, hence, shift the contents of HAMED.COM by 256-byte. In this case the "TMV=503" will be shifted from location 500 to location 756. Clearly, the condition will be not satisfied in location 500 and location 756. Conditions that link the content of LOC with the contents of its predecessor or successor locations are not affected by the virus infection. For example, the condition:[LOC]=[LOC-1]+[LOC+1].

SDR Module Standard Format

The SDR module consists of the SDR code/data and trapdoor marking information. Two problems must be considered before defining the standard format of the SDR module:

Pseudo Trapdoor Mark: "The pseudo trapdoor mark is any value satisfy the searching formula and cause LEP to transfer control to an incorrect location". The increased probability of pseudo trapdoor mark in the value-based technique is the great disadvantage of this technique. For example, in Tab.1 the TMV associated with HAMED.COM is 220 with respect to LEP1. Clearly, this value may exist more than one time within the code or data of HAMED.COM. Moreover, viruses may consider this as a vulnerable spot for attack. For example, TMV will range from 0 to 255 if represented as one byte storage location. The virus can insert a table that contains all of the expected values (from 0 to 255) in the start of the infected executable. When LEP search the infected executable, it will find a pseudo trapdoor mark in the virus table and transfer control accordingly. This transfer will activate the virus to start its execution.

The probability of finding a pseudo trapdoor mark in the condition-based technique is less relative to the value-based technique. Because it searches for a **value that satisfy a condition.** For example, in Tab.2, even though the value 503 may exist many times within the code or data of HAMED.COM, it will be considered as a trapdoor mark only if it exists at offset address 500.

In general, the solution to the pseudo trapdoor mark problem is to use two TMVs associated with special integrity check information. Tab.3 shows the suggested standard format of the SDR module. According to Tab.3, LEP can find the SDR module as follows:

1. **P=0**

2. **IF (f11(P,LBC11)=g11(P,LBC12)) THEN 6**
3. **P=P+1**
4. **IF P<Executable_Size THEN go to 2**

;Trapdoor not found,

5. **Display Alarm Message "Cannot Find the Trapdoor".**

;Test if the second condition satisfied

6. **IF (f12(P,LBC13)=g12(P,LBC13)) THEN go to 8**

;TMV2 Not exist, continue the search

7. **Go To 3**

;Perform checksum test

8. **LEN= $f13^{-1}$([P+08H])**
9. **Using CHKS-ALG1, Set C= Checksum of the block (P+10H+LEN)**
10. **IF C=[P+0CH] THEN go to 12**

;Checksum Error, Continue The search

11. **Go To 3**

;Trapdoor Mark Found

12. **LOC1=P**

Destruction: Destruction can be caused by a virus infection. If the virus, for example, inserts part or all of its code between TMV11 and TMV12, between the header and the SDR module, or inside the SDR module. The solution to this problem is to use redundancy. That is, two SDR modules must be used so that if one of them destroyed the other one can be used. The first SDR module (shown in Tab.3) will be called the **"Primary SDR Module"** and the second SDR module will be called the **"Redundant SDR Module"**. The redundant SDR module is shown in Tab.4.

LEP-To-LEP Communication

As mentioned earlier, each LEP uses its own standard to generate and search for TMV and LOC. Users may ask what happens when the protected executable 'E' copied from the environment of LEP1 to the environment of LEP2. That is: **"How LEP2 can find LOC1 and LOC2 of 'E', without giving viruses the opportunity to find them".** Clearly, there must be a standard and secure communication protocol between LEP1 and LEP2. There are many ways to do this.

First, the communication can be done by using a global TMV, say "HMS". When the user asks LEP1 to copy 'E' to be used by LEP2, LEP1 must store "HMS" at LOC1 and LOC2. When the user asks LEP2 to execute 'E' for the first time, it will search 'E' for the global TMV (i.e. HMS) to find LOC1 and LOC2. Once "HMS" is found LEP2 will recalculate LOC1, LOC2 and reorganize the SDR locations according to its formula. The advantage of this method is its transparency for users. The great disadvantage of this method is its vulnerability. Viruses know that the global TMV used in copies is "HMS", therefore, they can easily find LOC1 and LOC2 of the protected executable.

Second, the communication can be done through the computer user. The user can ask LEP1 about the value of LOC1 and LOC2 of 'E' after copying the executable to a new diskette. The user must give the value of LOC1 and LOC2 to LEP2 (i.e. manually) when he wants to execute 'E' for the first time. The advantage of this method is that it is secure, because there is no way that a virus will know the value of LOC1 and LOC2. The disadvantage of this method is that it depends on the computer user responsibility. A serious problem arises when the user decides to copy a large number of executables. For example, if the user wants to copy 100-executable from DISK1 to DISK2, then, he must ask LEP1 about the value of LOC1 and LOC2 for each one of these executable, keep these (200-Value) values in his mind, and give them to LEP2.

Third, the best suitable method is to use Message-Files. When the user asks LEP1 to copy the 100-executable (E1, E2,… E100) from DISK1 to DISK2, LEP1 will ask the user about

the message file name to store the values of LOC1 and LOC2 in it. Assuming that the user uses the name MESSF.LOC. LEP will create a message file with the name MESSF.LOC in DISK2, and insert 100-entry in this file. Each entry represent one executable. The standard entry format is shown in Tab.5. LEP2 can use the message file to find LOC1 and LOC2 of each executable individually or all executables at once. Regardless of the number of executables being copied, the user needs to remember only the message file name.

-The Communication Standards

Before describing the communication protocols used in PHASE-ONE and PHASE-TWO; and the operations of LEP and SDR in each phase, the standard data blocks, variables, and flags used in the system must be defined and described. Fig.2 shows an overview of the communication system and its individual components. Each data block, variable, and flag used for specific purpose. The following sections will describe these components and explain the purpose of using them.

- Image Information Block (IIB)

In PHASE-ONE, LEP must prepare the Image Information Block (IIB) before transferring control to the SDR. The IIB describes the status of the disk and memory images of E. The standard format of the IIB is shown in Tab.6.

MIS=DIS for COM and SYS files. But MIS≠DIS for EXE and OVL files, because the header exist in the disk image but not loaded with the memory image. MIS and DIS are used by the detection algorithm. TOA is used by the detection and eradication algorithms as will be explained later. EPN (Executable Private Number) is generated by LEP for each executable before executing it. EPN of E is considered, by LEP, as the “identifier” or “Secret Name” of E. While in PHASE-TWO, the SDR of E must pass the EPN of E associated with the other identification information to LEP, so that it can get the permission from LEP to access the disk image of E. LEP-TRAP is a pointer to a special trapdoor within LEP. This trapdoor is used to ensure that the SDR can communicate with LEP secretly while in PHASE-TWO.

- Detection and Eradication Information

The SDR must know the following information about the protected executable:

1- **Executable Critical Bytes: “The critical byte is any byte which is virtually guaranteed to be changed after a virus infection”.** The executable size, the first three bytes of a COM or SYS, and the EXE or OVL file header are considered as critical bytes. Determining the critical bytes of a given executable, require a deep understanding of how the computer virus infect it. In general, the following components are calculated and stored in the SDR of E during the protection process “see section 5”:
   1. **CMIS= Correct Memory Image Size (CMIS)**
   2. **CDIS= Correct Disk Image Size (CDIS)**
   3. **The correct values of any other critical byte. Such as the first three bytes of COM and SYS files and the header of EXE and OVL files.**

- **Integrity Check Information (Checksum):** The SDR is assumed to view the executable memory image as group of N-Blocks (BLK1,BLK2, ... BLKN) “see Fig.3”. The block size (BLKS) is equal for all blocks and stored in (BLKS). A CheckSum Number (CSN) calculated for each block using the algorithm CHKS-ALG, and stored in a special SDR table. Also, a given block ‘BLKi” is assumed to be valid (i.e. has a correct checksum) if the following condition satisfied: **“CHKS-ALG(BLKi, CSNi)=0”.**
- **Position Test Information:** The position test information (SOF and EOF) are used (with TOA, and MIS) by the FBEA to find the virus block position relative to the SDR module as show in Tab.7. Fig.3 shows how the protected executable appear in memory. SOF and EOF represent the position test information and defined as follow:

- **The Start Offset (SOF):** Is the offset of the SDR trapdoor relative to the start of the memory image**.**
- **The End Offset (EOF):** Is the backward offset of the SDR trapdoor relative to the end of memory image**.**

- SDR Internal Flags

SDR uses two internal flags:

1- Executable Infection Flag (EIF): SDR set this flag if it detects a virus infection. The status of this flag is returned to LEP at the end of PHASE-ONE.

2- Executable Repair Flag (ERF): SDR set this flag if it can repair the infected executable. The status of this flag is returned to LEP at the end of PHASE-ONE, and the end of PHASE-TWO through LEP-TRAP.

2.4 The Secret Identification Block (SIB)

The SIB passed by the SDR of E to LEP during PHASE-ONE. When LEP switch to PHASE-TWO, it can use SIB to distinguish the SDR of E from the SDRs of the other executables and viruses. The standard format of SIB is described in Tab.8.

3- PHASE-ONE Communication

PHASE-ONE starts when LEP receives a request to execute an executable. Assuming that LEP receives a request to execute the executable E, the following sections describe the sequence of operations:

Locating and Executing SDR

The following steps describe loading the executable by LEP and preparing the IIB:

Step-1: "Loading the executable memory image"

1. **Find the disk image of E.**
2. **Store DIS in the IIB.**
3. **Assign EPN to E and store it in the IIB.**
4. **Store LEP-TRAP in the IIB**
5. **Reserve a memory block to store the executable memory image. And store the start address (SMI) in the IIB.**
6. **Load the executable memory image into the reserved block.**
7. **Store the MIS in the IIB.**

Step-2: "Finding the Trapdoor"

1. **Search for the trapdoor of the primary SDR module.**
2. **IF (the trapdoor found) THEN go to 6**
3. **Search for the trapdoor of the redundant SDR module.**
4. **IF (the trapdoor found) THEN go to 6**
5. **LEP cannot find any one of the trapdoors. This can happen if the executable is not Self-Protected, the executable is new in the system, or both SDR modules are destroyed due to a virus infection. In either case, LEP must display alarm message to the user, and ask him what to do. Depending on the user response, LEP must proceed as follow:**
   - **If E is not self-protected, then execute it in a virtual computer (explained later).**
   - **If E is new, then prepare it using the message file.**
   - **If E destroyed by a virus infection, then, avoid executing it**
6. **Store the offset of the trapdoor in TOA.**
7. **Store TOA in the IIB.**

Step-3:"Execute and Wait"

LEP can transfer control to the SDR trapdoor using a far call instruction, and wait until the SDR return control again. What happens when the SDR receives control is explained in the next section.

SDR Operation in PHASE-ONE

The SDR operation can be described by two cycles: **"Self-Test Cycle, and Self-Repair Cycle"**. Both cycles are described in the following sections.

Self-Test Cycle

This cycle initiated each time SDR executed, in this cycle SDR must decide whether E is infected or not. If E is infected, SDR will set EIF, and clear it otherwise. The status of EIF returned from SDR to LEP as shown in Fig.2. The following algorithm describe the self-test cycle:

;Size test
IF DIS≠CDIS THEN go to 13
1. IF MIS≠CMIS THEN go to 13
;Critical byte test
2. IF (Any critical byte changed) THEN go to 13
;Checksum test
3. i=1
4. C=CHKS-ALG(BLKi, CSNi)
5. IF C≠0 THEN go to 13
6. i=i+1
7. IF i ≤ N THEN go to 5
;Position test
8. IF SOF≠TOA THEN go to 13
9. IF EOF≠MIS-TOA go to 13
"The executable is clean".
10.EIF=0
11.Return to LEP.
"The executable is infected".
12.EIF=1
13.Go to the Self-Repair Cycle.

Self-Repair Cycle

In this cycle, SDR will try to eradicate the virus and repair the protected executable. SDR will set ERF if the protected executable repaired properly and clear it otherwise. A new eradication algorithm (Foreign Block Eradication Algorithm "FBEA") will be used. FBEA capability depends on how the virus distributes itself within the infected executable. The FBEA described below can eradicate SBD viruses efficiently, and can be upgraded to eradicate CBD viruses as well. However, because the CBD idea not used by viruses yet "see [3]", the discussion will be limited for the SBD viruses only. Assuming a simple virus distribution, the eradication algorithm can be divided into the following steps:

Step-1: "Find the virus block position relative to the SDR"

In general, if the virus block inserted after SDR, the infected executable memory image will take the form of image ①,②, or ③ in Fig.4. Therefore, the search must start from BLK1. If the virus block inserted before SDR, the infected executable memory image will take the form of image ④,⑤, or ⑥. Therefore, the search must start from BLKN. Finding the virus block position relative to SDR can be done by using the position test information TOA, MIS, SOF, and EOF as shown in Tab.7.

Step-2: "Starting the block checksum test"

For example, let us assume that the virus is inserted after the SDR, therefore, the block checksum test must start from BLK1. The following algorithm can be used in this case:

1. **i=1**
2. **IF BLKi contain any critical byte, then repair the critical bytes.**
   "Only BLK1 in image ①, ②, and ③ affected by this step"
3. **C=CHKS-ALG(BLKi, CSNi) ;Calculate the checksum of BLKi**
4. **IF C≠0 THEN go to Step-3 ;Invalid block**
5. **Move BLKi into BUF ;Valid block**
6. **i=i+1**
7. **IF i ≤ N THEN go to 2**

Arriving to this point means that the virus has inserted all of its added bytes at the end of the infected executable, as shown in image ①.

Go to Step-6

"Note the difference between (go to 3) and (go to Step-3)"

Step-3: "Reversing the block checksum test order".

Arriving to this step means that the virus inserts its added bytes after SDR but not at the end of the infected executable. In this case, the infected executable memory image expected to take the form of image ② or ③. Because BLKi is not found, the searching processes must be reversed. The following algorithm can be used:

1. **j=0**
2. **m=N-j**
3. **IF BLKm contain any critical byte, then, repair the critical bytes.**
4. **C=CHKS-ALG(BLKm, CSNm) ;Calculate checksum number of BLKm**
5. **IF C≠0 THEN go to Step-4 ;Invalid checksum**
6. **Move BLKm into BUF ;Valid checksum**
7. **IF m=i THEN go to 10**
8. **j=j+1**
9. **Go to 2**

Arriving to this point means that the virus block inserted between two consecutive blocks (BLKi-1 and BLKi) without destroying any one of them, as shown in image ③. Because the blocks (BLK1, BLK2, ..., BLKi-1) are moved into BUF in Step-2. And the blocks (BLKi, BLKi+1, ... BLKN) are moved into BUF in Step-3. Therefore, all blocks are moved into BUF.

10. **Go to Step-6**

Step-4: "Repairing the damaged block"

Arriving to this point means that the virus has inserted its block inside BLKi and, hence, destroyed this block. This is shown in image ②. Two routines to repair the destroyed block BLKi will be discussed:

REPAIR1:

Inputs:

- **i= Block number**
- **S= Start offset address of BLKi which contain the virus block.**
- **E= End offset address of BLKi which contain the virus block.**
- **Z= MIS-CMIS=NAB**
- **P= Offset address of a byte that is guaranteed to exist in the virus block.**

Outputs:

- **C=0 ;BLKi cannot be repaired**
- **C=1 ;BLKi repaired properly and stored in the buffer BBK**

The Idea: REPAIR1 assumes that the virus block start at P and define: **"VS (Virus-Start)=P, and VE (Virus-End)= VS+Z-1".** And then, moves the bytes at block (S-To-(VS-1)) and block ((VE+1)-To-E) into the buffer BBK. If the checksum test on BBK fails, REPAIR1 assumes that P is not the actual start of the virus block. That is, there is at least one byte belong to the virus block and exist before P. Therefore, REPAIR1 shift VS and VE up by one byte position, and repeat the test process.

Algorithm:

1. **VS=P**
2. **VE=VS+Z-1**
3. **IF {VE>(E-1)} THEN {[VS=VS-(VE-(E-1))] AND [VE=VS+Z-1]}**

   "This ensures that VE≤(E-1). Note that, the byte at E must belong to BLKi, because, otherwise, all of the content of BLKi exists above the virus block. And this can happen only if BLKi wasn't destroyed by the virus infection".
4. **Move the bytes at block (S-To-(VS-1)) into BBK**
5. **Move the bytes at block ((VE+1)-To-E) into BBK**
6. **C=CHKS-ALG(BBK, CSNi)**
7. **IF C= 0 THEN go to 12 ;Valid block**

   "Arriving to this point means that VS is not the actual start of the virus. That is, there is at least one byte belong to the virus but exist above VS. Therefore, VS must be decremented and the checksum must be calculated again".
8. **VS=VS-1**
9. **VE=VS+Z-1**
10. **IF VS>S THEN go to 4**
11. **C=0, and RETURN "BLKi cannot be repaired"**
12. **C=1, and RETURN "BLKi repaired properly and stored in BBK"**

REPAIR2:

Inputs:

- **i= Block number**
- **S= Start offset address of BLKi which contain the virus block.**
- **E= End offset address of BLKi which contain the virus block.**
- **Z= MIS-CMIS=NAB**

Outputs:

- **C=0 ;BLKi cannot be repaired**
- **C=1 ;BLKi repaired properly and stored in the buffer BBK**

The Idea: REPAIR2 is based on the fact: **"If BLKi *destroyed* by one sequential virus block, then, the byte at S and the byte and E must belong to BLKi. Because otherwise, if the byte at S(E) belong to the virus block, the entire virus block must exist above (below) BLKi, and this can happen only if BLKi wasn't *destroyed* by the virus block insertion".** REPAIR2 moves one byte starting from S and (BLKS-1) bytes starting from E into BBK. If the checksum test on BBK fail, REPAIR2 repeat the process by moving 2-bytes starting from S and (BLKS-2) bytes starting from E into BBK.

Algorithm:

1. **k=S**
2. **j=k+Z+1**
3. **Move the bytes at block (S-To-k) into BBK.**
4. **Move the bytes at block (j-To-E) into BBK.**
5. **C=CHKS-ALG(BBK, CSNi)**
6. **IF C=0 THEN go to 10;Valid block**
7. **k=k+1**
8. **IF k< (BLKS-1) THEN go to 2**

9. **C=0, and RETURN "BLKi cannot be repaired"**
10. **C=1, and RETURN "BLKi repaired properly and stored in BBK"**

The advantage of REPAIR2 over REPAIR1 is that it can work without using P. However, the number of trails or the time needed by REPAIR1 is, in general, less than the time needed by REPAIR2. Which routine "REPAIR1 or REPAIR2" the SDR must use, depend on the prepared input arguments. The input arguments can be prepared as follow:

1. **S= (i-1)*(BLKS) ;The number of bytes in all blocks before BLKi.**
2. **E= S+Z+BLKS;To understand how S and E calculated see Fig.5.**
3. **The SDR can use REPAIR1 only if it can prepare P. Otherwise, it must use REPAIR2. Two ways are suggested to determine P:**
   a) **The entry point of the virus code within BLKi can be found from the standard entry point of the executable. For example, for COM files the displacement of the near jump instruction which is found in the first three bytes of BLK1 can be used to find the virus entry point. After finding this entry point, set P= Virus code entry point.**
   b) **If the virus code entry point cannot be found, the following fact can be used: "IF (Z>BLKS) THEN (The point at "S+(E-S)/2" must be in the virus block)". Therefore: IF (Z>BLKS) THEN (P=S+(E-S)/2)".**

If P prepared, the SDR can proceed as follows:

1. **CALL REPAIR1**
2. **If C=1 THEN go to 6 ;BLKi repaired properly using REPAIR1**
3. **CALL REPAIR2**
4. **IF C=1 THEN go to 6 ;BLKi repaired properly using REPAIR2**

BLKi cannot be repaired using REPAIR1 or REPAIR2.

5. **Go to Step-5**
6. **Move the contents of BBK into the gap of BLKi in BUF.**
7. **Go to Step-6**

Step-5: "Return Error Code and SIB to LEP".

Arriving to this step means that the SDR cannot repair E properly. Therefore, the SDR must return the following information:

1. **EIF=1 ;The file is infected**
2. **ERF=0 ;The file is not repaired**
3. **SIB**
4. **Go to Step-7.**

Step-6: "Repair the Infected Disk Image of E"

Arriving to this step means that BUF contain all blocks (BLK1, ... BLKN) of the protected executable. The content of BUF represent a clean memory image of the executable, therefore, it can be used to repair the infected disk image of E. The following algorithm describes how this can be done.

1. **If the infected executable is an EXE or OVL file, then, put the correct file header in the start of BUF.**
2. **Overwrite the disk image of E by the content of BUF.**
3. **ERF=1 ; Executable Repaired Properly.**
4. **Go to Step-7**

Step-7: "Return To LEP"

The SDR can return control to LEP now. PHASE-ONE terminated at this point.

- PHASE-TWO Communication

LEP starts PHASE-TWO after the SDR execution, in PHASE-ONE, is completed. Once in PHASE-TWO, LEP will test the status of EIF and ERF returned from SDR. The executable is clean if (EIF=0). Therefore, LEP can execute the executable through its standard entry point. The executable was infected but repaired properly by SDR if (EIF=1 and ERF=1). Therefore, LEP must load the executable disk image again. And, then execute it through its standard entry point.

The executable is infected and SDR cannot repair it if (EIF=1 and ERF=0). Therefore, the Virus Follower Eradication Algorithm (VFEA) must be used to eradicate the virus. VFEA is based on the fact that shell type viruses, "See [7]", always repair the infected executable memory image before executing it. Therefore, VFEA use the following eradication approach: **"The infected executable must be executed through its standard entry point so that the attached virus will get the opportunity to execute. The virus will repair the memory image of the executable and execute it. SDR will receive control again and test the memory image. If the memory image repaired properly by the virus, SDR must alert LEP to use the current memory image of the executable to replace the infected disk image of the executable".**

However, giving the opportunity to the virus to execute is very critical. Therefore, before executing the virus, LEP must prepare a trusted environments "Virtual Computer System" to ensure that the virus cannot cause any damage or infection. Preparing the virtual computer system and executing the SDR in PHASE-TWO is described in the following sections:

The Virtual Computer System

The idea of the virtual computer system is described in [4]. In general, the virtual computer system must be designed to satisfy the following requirements:

1- **Protecting System Disks:** Prevent the virus from infecting or destroying any target site/cell in the system disks.
2- **Protecting System Memory:** Prevent the virus from reserving memory space and hide itself there, that is, stay resident in system memory.
3- **Deceiving the Virus:** The virtual computer system must give the virus the illusion that it is running in a normal system. Because, if the virus knows that it exists in a virtual computer system, it may try to use special methods to bypass or deceive the virtual computer system; or it may terminate its execution without repairing its host, therefore, VFEA cannot repair the protected executable.

Satisfying these requirements depend on the computer system hardware (i.e. Real-Mode or Protected-Mode PC) and software (i.e. DOS or WINDOWS). As a case study, preparing a virtual computer system in DOS machines will be explained in what follows:

1- **Protecting System Disks:** In order to prevent the virus from infecting/destroying target sites/cells in the system disks, it must be prevented from writing to these disks. First of all, LEP must ensure that all of the disk related interrupt vectors points to special handlers. Therefore, LEP must redirect the following vectors:
   1. **Vector 13H (BIOS Disk Interrupt INT 13H).**
   2. **Vector 21H (DOS-API Interrupt INT 21H).**
   3. **Vector 25H (DOS: Absolute Disk Read Interrupt INT 25H) "see [5]".**
   4. **Vector 26H (DOS: Absolute Disk Write Interrupt INT 26H).**

   Under DOS, if the virus knows that it exists in a virtual computer system, it may try to use direct hardware access method to bypass the virtual computer handlers. The virtual computer can use one of the following methods to prevent viruses from accessing the disk: First, the handler of the virtual computer can reject any write to disk request by returning some error code which indicates that the requested operation

cannot be performed because, for example, the drive is not ready. Rejecting all writes to disk requests can help the virus in deciding whether it is working in a virtual computer system or not. Second, the virtual computer system can trick the virus to believe that the requested write to disk operation is performed while it is not. This can be done by returning a no error code that indicates the requested operation performed properly, without performing the actual operation. If this method used, the virus can decide whether it is working in a virtual computer system or not by using the following trick:

1. **Send a request to write the data block (BLK) to disk (C:).**
2. **Read the data block from disk (C:) into (BLK1)**
3. **If BLK≠BLK1, then, the data wasn't written to the disk. Therefore, the system is a virtual computer system.**

Finally, all disk read/write operation can be redirected to a special RAM disk instead of the actual disk. In this way, the virtual computer system can trick the virus to believe that the requested write to disk operation was performed and the data written to the disk properly, while the data was written to the RAM disk. If the virus request the data later, the virtual computer can read it from the RAM disk. This means that the virtual computer must handle both disk read and write operations, this explains why vector 25H was redirect above.

2- Protecting the IVT: The virtual computer system must save the content of the IVT before executing the virus, and restore it once the virus execution completed. This ensures that any redirection to the IVT vectors by a resident type virus is fixed once the virus execution is completed.

3- Protecting System RAM: Protecting the PC system RAM can be done as follow:

- **Save the amount of the available URAM (the word at 0:0413H) before executing the virus. And restore the content of 0:0413H once the virus execution completed. This ensures that the virus cannot decrement the amount of URAM and install itself at the end of the URAM.**
- **If the virus tries to allocate a memory block using function 48H of INT 21H, then, the virtual computer must store the address of the allocated memory block so that it can release this block once the virus execution completed. Therefore, it must redirect vector 21H.**
- **Save a map of all memory control blocks "see "[٥] before executing the virus. And restore them once the virus execution completed. This ensures that the virus cannot allocate memory by directly accessing the memory control blocks.**
- **Redirect vector 66H to special handler before executing the virus. INT 66H is used to access the Expanded Memory Manger functions. This interrupt must be handled only if the system uses expanded memory. If the virus allocate a page (or pages) in the expanded memory, the handler must store the handles reserved by the expanded memory system for the allocated pages. This handle can be used to release the reserved pages when the virus execution completed. Expanded memory expands RAM beyond the 640KB limit, for more information see [5].**
- **Extended memory exists in AT machines (the memory beyond the 1 MB limit "see [5]"). Because the eXtended Memory Manager functions are called through a FAR CALL instruction, instead of the special interrupt, the virtual computer cannot intercept requests to the extended memory function. Therefore, the virtual computer must ensure that all of the extended memory is free before executing the virus. And free it again after the virus execution completed, to ensure that viruses cannot allocate memory in the extended memory area.**

Execute SDR and Wait

After preparing the virtual computer system, LEP can execute E through its standard entry point and wait until it receives a special call which indicate that the virus execution was completed and SDR was executed. Note that LEP can switch from the virtual computer system back to the normal system only after ensuring that the virus execution was completed. Receiving the call is the clue that indicates to LEP that SDR is the currently active program. Because viruses always try to disable or circumvent the protection mechanism, the following requirements must be satisfied:

1- **The virus cannot mask the call:** Therefore, in the PC system the software interrupt mechanism cannot be used to perform the call. Because the virus can redirect all of the interrupt vectors to its own handler, therefore, it can mask the call and preventing it from arriving to LEP. The call must be direct, that is, by using jump or call instructions. This explains why LEP passes a pointer (LEP-TRAP) to its trapdoor to SDR in PHASE-ONE.

2- **The virus cannot deceive LEP:** In order to deceive LEP by a tricky call, the virus must know where to send the call and which information to pass with it. Therefore, the virus cannot deceive LEP through LEP-TRAP, because it doesn't know the correct values of LEP-TRAP, EPN, and SIB. Since these variables passed from LEP to SDR in PHASE-ONE and the virus was inactive at that time, therefore, there is no way to know these variables by the virus.

SDR Operation in PHASE-TWO

When the SDR receives control in PHASE-TWO it will proceed as follows:

1. **IF EIF=1 THEN go to 3**
   ;E is clean
2. **Continue the execution of E normally**
   ;E is infected, start the repair cycle using VFEA
   ;Perform checksum test
3. **i=1**
4. **C=CHKS-ALG(BLKi, CSNi)**
5. **IF C≠0 THEN go to 12**
6. **i=i+1**
7. **IF i≤N THEN go to 4**
   ;All blocks are valid.
8. **ERF=1 ;The memory image is repaired properly**
9. **SMI= Start of the repaired Memory Image**
10. **MIS= Size of the repaired Memory Image**
11. **Go to 15**
    ;BLKi is invalid
12. **ERF=0 ;The memory image cannot be repaired**
13. **Prepare SIB**
14. **EPN= Executable Private Number that was received from LEP in PHASE-ONE**
    ;Transfer control to LEP through LEP-TRAP
15. **JMP FAR PTR LEP-TRAP**

Receiving Control from SDR

The virtual computer system prevents any program from writing to the system disk to ensure that viruses cannot replicate themselves. However, the question is how the SDR of E can repair the infected disk image of E. LEP can give the permission to the SDR to perform disk write operations. SDR must call LEP through LEP-TRAP with the proper EPN and SIB to get this permission. Giving this permission to SDR can be done in one of two ways:

1- **Filtering:** In this case, the virtual computer must be capable of distinguishing between the SDR write to disk requests and the other requests. This can be done, for example, by sending EPN and SIB with each request.

2- **Restoring the IVT:** LEP can restore the IVT content when receiving EPN and SIB through its trapdoor. Therefore, SDR can access the disk as desired by using DOS and BIOS services.

However, giving the permission to SDR to access the disk is not a good idea, because, the virus exist in system memory and may, in some way, interfere with the SDR operation, so there is a possibility that the virus will get the opportunity to access the disk. Therefore, the following method is suggested: **"Instead of giving the disk access permission to SDR, LEP can repair the disk image by itself after receiving the necessary information from SDR".**

Therefore, as shown above, SDR return SMI and MIS to LEP through LEP-TRAP. After receiving this information, LEP can proceed as follows:

1. **Compare the received EPN with the one that was given to SDR in PHASE-ONE, if not equal go to 6**
2. **Compare the received SIB with the one that was received from SDR at the end of PHASE-ONE, if not equal go to 6**
3. **IF ERF=0, THEN, switch to normal mode "SDR cannot prepare a clean backup image".**

   "A clean backup image is available and LEP must use it to replace the infected disk image, as follows:"
4. **Use the MIS-byte block that starts at (SMI) to replace the executable whose private number is EPN.**
5. **Switch to normal mode.**
6. **The SDR identification information (EPN or SIB) are invalid, therefore, reject the call, and display alarm message.**

4.5 Switching to Normal Mode

LEP can switch to normal mode after receiving the SDR request through LEP-TRAP and repairing the infected disk image if necessary. LEP must do the following so that it can switch to normal mode:

1. **Restore the IVT**
2. **Release any memory block allocated during the virus execution using DOS function 48H, URAM, memory control blocks, expanded memory, and extended memory.**
3. **Release the memory allocated for E, to ensure that the viral code which exist at this memory area will be destroyed.**
4. **Release the RAM disk memory area.**

5- Vaccination

There is a large number of executables in the world. All of these executables are developed without any built-in SDR. Clearly, some way must be found to convert these executables to self-protected executable. Vaccination can be used to do this. In this research, vaccination is defined as: **"The process of converting an executable into a self-protected executable by injecting the SDR module inside it".**

"**Note:** Some references (see [6]) refer to 'Vaccination' as one of the protection methods that was suggested by IBM. Some anti-virus programs are designed to inject themselves inside the executable files, and then operate like a '**Benign Virus**' when the vaccinated program executed. The injected anti-virus creates a signature (finger-print) of uninfected executable, and display an alarm if the signature or executable size changed. Unfortunately, the alarm generated after the virus execution. The difference between the suggested SDR and the Vaccine of IBM, is that the SDR executed before any attached virus and uses sophisticated techniques to detect and eradicate the virus".

In fact, the idea of vaccination is borrowed from viruses. A special program "called the INJECTOR" will inject the executable by the vaccine (i.e. SDR). However, there is no searching, or infection routine in this vaccine, therefore, the vaccine cannot replicate itself. The word "injection/inject" is used instead of the word "infection/infect" to distinguish vaccination from virus infection. The following steps can describe the algorithm used by INJECTOR to inject SDR into an executable E.

Step-1: "Injecting SDR in the executable E"

This is similar to what the virus infection routines do, that is:

1. **Store the contents of the standard entry point of E in SDR.**
2. **Attach SDR to E.**
3. **Redirect the standard entry point of E to point to the entry point of SDR.**

After this injection, E will take the form of image ② in Fig.6. From now, image ② is considered the correct executable image. If SDR receives control through the standard entry point and after the completion of its execution, it can repair the image of E to take the form of image ① so that it can execute properly.

Step-2: "Preparing the SDR Module"

As shown in Fig.6, the correct image of E is image ② from the SDR point of view. Therefore, INJECTOR must prepare SDR with respect to this image. After calculating and storing SOF, EOF, CDIS, CMIS, and the other critical bytes, INJECTOR must divide image ② into block (BLK1, ... BLKN) and store the checksum numbers in the SDR table see fig. 3. Preparing the SDR header can be done easily if INJECTOR knows which algorithms and functions are used by LEP. Otherwise, INJECTOR must store the values of LOC1 and LOC2 in a message file so that LEP can prepare the SDR module header.

- Efficiency Analysis

The efficiency of SDS must be analyzed to see whether it can satisfy all of the requirements of IAVS or not. The efficiency of the SDS is determined according to the analysis criteria presented in [٤]. In what follows, SDS will be analyzed with respect to detection, prevention, eradication, and damage control.

1- Detection: SDR classified as a Target-Based detection.

a- VGS= Max.: Because SDR knows every thing about the protected executable, it can detect any change to this executable. Viruses cannot avoid detection by the SDR, because, they cannot infect the executable without changing its contents. Even stealth type viruses "see [7]" cannot deceive the SDR, because, they cannot get the opportunity to execute before the SDR. Therefore, SDR detection capability is independent from the virus generation date and VDC. Also, if LEP cannot find any one of SDR modules, it concludes that the protected executable was changed, and it might be infected by a virus.

b- TGS=1: Because SDR designed to detect viruses attached to the protected executable only.

c- Capability: The following is concluded from a & b: **"SDS can satisfy the requirement of ideal Target-Based detection".**

- Prevention: Both VEP-Based and VIP-Based prevention can be used in SDS.

- **VEP-Based Prevention:** SDR classified as a VEP-Based prevention anti-virus, because, it can detect the infection without giving viruses the opportunity to execute.

  a- VGS= Max.: SDR prevention is independent from the virus generation date and VDC.

  b- TGS= Max.: SDR will prevent the detected virus from infecting any target site within its environment.

c- Capability: The following is concluded from a, b, and the fact that the SDR is ideal with respect to detection: **"SDS can satisfy the requirement of ideal VEP-Based prevention".**

- **VIP-Based Prevention:** LEP classified as a VIP-Based prevention anti-virus, because, it can prevent the detected virus from infecting a new target site when executed in PHASE-TWO by using the virtual computer system.
  - **a- VGS:** It was mentioned in section 4.1 that the direct hardware access is the most vulnerable spot that the virus can use to penetrate the virtual computer system. Maximizing VGS thoroughly depends on the system that will implement the SDS. VGS can be maximized, if SDS is implemented on a computer system that provides hardware protection level (i.e. protected mode PC) with a permission or privilege level that will prevent an unauthorized program to direct access the hardware resources. Otherwise, VGS cannot be maximized.
  - **B- TGS= Max.:** LEP will try to prevent the detected virus from infecting any target site within its environment.
  - **c- Capability:** The following is concluded from a & b:
    - **"SDS can satisfy the requirement of ideal VIP-Based prevention if used on a system that provides hardware protection level (i.e. Protected Mode PC)".**
    - "**SDS cannot satisfy the requirement of ideal VIP-Based prevention if used on a system that lacks the hardware protection level (i.e. Real Mode PC)".**

- Eradication: As shown above, two eradication algorithms (FBEA and VFEA) are used by SDS:

- **FBEA:** FBEA uses Analysis-Based eradication technique. It analyzes the protected executable and eradicates any foreign block.
  - **a- VGS:** FBEA eradication capability depend on the virus type, shell or intrusive "see [7]", and NAB-distribution method (SBD or CBD). Therefore, the FBEA described above can eradicate only shell type viruses that use SBD method. Even though, FBEA can be upgraded to eradicate CBD viruses, the virus designers can design new strains of CBD-viruses that cannot be eradicated by FBEA. Therefore, the FBEA eradication capability depends on the virus generate date. This means that: **"FBEA cannot maximize VGS".**

NOTE: FBEA is a good enhancement in the design of eradication algorithms. It is very efficient eradication algorithm at the present time, because, it can eradicate, almost, all of the currently available viruses.

  - **b- TGS=1:** Because FBEA designed to eradicate viruses from the protected executable only.
  - **c- Capability:** The following is concluded from a & b: **"SDS cannot satisfy the requirement of ideal Analysis-Based eradication".**
- **VFEA:** VFEA uses Backup-Based eradication technique. In this case, the infected disk image represents the vulnerable copy, and the repaired (i.e. by the virus) memory image represents the backup copy.
  - **a- VGS:** VFEA can eradicate shell type viruses only and its eradication capability depend on the virus generation date, this means that: **"VFEA cannot maximize VGS".**
  - **b- TGS=1:** VFEA designed to eradicate viruses from the protected executable only.
  - **c- Capability:** The following is concluded from a & b: **"SDS cannot satisfy the requirement of ideal Backup-Based eradication".**

NOTE: VFEA is a good enhancement in the design of eradication algorithms. It is very efficient eradication algorithm at the present time, because, it can eradicate, almost, all of the currently available shell-type viruses.

- **Damage Control:** SDS system can use Virus-Based, Cell-Based, and Backup-Based damage control techniques.
  - **Virus-Based Damage Control:** In PHASE-ONE, SDR classified as a Virus-Based damage control anti-virus. Because it detects the virus before executing it. Clearly, the virus cannot cause any damage if not executed. Virus-Based damage control is similar to VEP-Based prevention, therefore, VGS= Max, TGS= Max, and: "**SDS can satisfy the requirement of ideal Virus-Based damage control".**
  - **Cell-Based Damage Control:** In PHASE-TWO, LEP classified as a Cell-Based damage control anti-virus, because, it tries to prevent viruses from destroying any target cell within its environment by using the virtual computer system. Clearly, direct hardware access is still a vulnerable spot. Cell-Based damage control is similar to VIP-Based prevention, therefore:
    - **"SDS can satisfy the requirement of ideal Cell-Based damage control, if used on a system with hardware protection level".**
    - **"SDS cannot satisfy the requirement of ideal Cell-Based damage control, if used on a system with no hardware protection level.**
  - **Backup-Based Damage Control:** In PHASE-TWO, LEP can store a backup copy of the virus target cells such as FAT, RD, BPS, and CMOS RAM before executing the virus. After the completion of the virus execution, LEP can repair any target cell that was destroyed by the virus, by using its backup. This technique is similar to saving the IVT before executing the virus and restoring it after the completion of the virus execution. In this case, LEP classified as a Backup-Based damage control anti-virus.

    a- VGS= Max.: LEP protection is independent from the virus generation date and VDC.

    b- TGS: TGS depend on the number of target cells selected by LEP.

    c- Capability: The following is concluded from a & b: LEP can protect any target cell in the selected T-group against destruction by any virus, therefore: **"SDS can satisfy the requirement of ideal Backup-Based damage control".**

- SDR Size Optimization

SDR will increase the protected executable size and the time needed to load and execute it. Therefore, minimizing the SDR size must be one of the design goals. One way to minimize the code of SDR is to implement the FBEA as a LEP (or OS) service. As mentioned earlier, the SDR calls the algorithm 'CHKS-ALG' with two arguments "Block number 'BLKi' and Checksum number 'CSN'". Clearly, it is not necessary to implement CHKS-ALG inside each SDR. Instead CHKS-ALG can be implemented as a global OS service that can be used by all SDRs. Note that the fact that the block size is different for different SDRs ensures that the general trend to avoid standard protection is not violated. Assuming that CHKS-ALG uses a standard CRC method, then, it can be implemented as a global service as follows:

CHKS-ALG

INPUT:

- BLKi= Start address of the block
- BLKS= Block Size in bytes
- CSN= The CRC checksum of the block.

OUTPUTS:

- C=0 ;Valid checksum
- C=1 ;Invalid checksum

In the same way, FBEA can be implemented as a global service. In this case, CHKS-ALG can be implemented inside the global FBEA. The standard Input/Output of the global FBEA can be defined as follows:

FBEA:

INPUTS:

- SMI, MIS, TOA, SOF, EOF
- N= Number of blocks in the memory image
- BLKS= Block Size
- Checksum Number Table: The table format is shown in Tab.9

OUTPUTS:

- C=0 ;Repair Succeed
- C=1 ;Repair Failed

- Conclusions and Discussion

A new approach in designing anti-virus system is presented in this paper. The proposed system is called Self-Defence System (SDS). The purpose of this research is to design an anti-virus system that complies with the requirements of ideal anti-virus system (IAVS) "see [4]. In [4], it was mentioned that Dependent-Defence System (DDS) couldn't satisfy the requirement of the IAVS because of four problems. Now let us consider these problems from the SDS point of view:

1- Lack of Knowledge: In principle, SDR must be implemented as integral part of the protected executable during the development process and it must know every thing about the protected executable. Therefore, SDS can satisfy the requirement of ideal detection. However, if the SDR injected inside an existing executable using vaccination, then, it is important to ensure that the executable is clean before injecting the SDR. If a virus exists in the executable prior to the injection, then, SDR cannot detect the presence of this virus.

2- Standard Protection: In principle, each SDR is designed and implemented by a different manufacturer. Therefore, the code and data of SDR will be different from one SDR to the other. For example, even though more than one SDR may use the FBEA, the block size (BLKS) selected by each SDR and the location of the checksum number table might be different. Clearly, viruses cannot deceive SDR without knowing this information.

3- Intended Vulnerability: In SDS, the SDR capability in coping with viruses is considered one of the executable quality factors. Good programs are those that can protect themselves against viruses efficiently. Therefore, any failure in the SDR operation and any undetectable infection have bad affects on the trustiness between the program developer and his customers. They may, simply, do not purchase his programs if they find that the SDRs associated with these programs are vulnerable.

4- Users Responsibility: In SDS, user responsibility is no longer a requirement or a factor of efficiency, because, the SDR designer cannot claim that his program cannot protect itself against the computer virus because the user do not use the protection program properly. SDR is embedded inside the protected program and its operation is, or must be, transparent for the computer user.

As a summary: **"SDS system solves the problems (Lack of knowledge, Standard Protection, Intended Vulnerability, and User Responsibility) found in DDS".**

Also, SDS provides a good support to design and use new sophisticated and efficient eradication algorithms. Even though FBEA is not ideal, it is considered a good enhancement in the design of Analysis-Based eradication programs, relative to the eradication method used in DDSs. In DDS, the eradication program can eradicate only 'Known' viruses. Any detected virus must be 'Identified' by security professionals, the identification information stored in a database, and a special eradication program designed to eradicate the virus. In SDS, FBEA eradicate any foreign block that exists between the protected executable blocks. FBEA doesn't care whether the foreign block belong to a known virus, unknown virus, logic bomb, time bomb, or a Trojan Horse.

Even though VFEA is not ideal, it is considered a good enhancement in the design of Backup-Based eradication programs, relative to the methods used by DDS. In DDS, for each target site (vulnerable copy), a backup copy must be stored in the disk. In SDS, VFEA depend on the virus to generate the backup-copy and then use it to repair the vulnerable copy, therefore, it has the following advantages relative to the method used by DDS:

- No disk space needed to store backups.
- There is no way that the virus will reside in the backup copy.
- There is no need to update the backup copy

References

-Garber, L., **"Antivirus Technology Offers New Cures"**, IEEE Computer Magazine, February, 1998.

-Pfleeger, C., P., **"SECURITY IN COMPUTING"**, Prentice-Hall Inc., 1989.

-Kaspersky, E., **"AVP Virus Encyclopedia"**, Version 1.3 (1992-1997).

-Hamid M. A. Abdul-Hussain, **"Computer Virology: Toward Designing An Ideal Anti-Virus System".** Engineering Journal, College of Engineering, University of Baghdad. Vol.8, No.3, 2002.

Tischer,M., **"PCINTERN SYSTEM PROGRAMING"**, Abacus, 1992

- عامر نزار فايز د. ،**"فيروسات الكمبيوتر"**، جامعة الإسراء-عمان-الأردن, 1995

- Hamid M. A. Abdul-Hussain & Hamed M. Shabib, **"Computer Virology: Formal Analysis of Computer Viruses".** Engineering Journal, College of Engineering, University of Baghdad. Vol.8, No.1, 2002.

List of Abbreviations

AT	Advanced Technology
BPS	Boot (or) Partition Sector
CBD	Complex Block Distribution
CDIS	Correct Disk Image Size
CMIS	Correct Memory Image Size
CRC	Cyclical Redundancy Check
CSN	CheckSum Number
DDS	Dependent Defense System
DIS	Disk Image Size
DOS	Disk Operating System
EBV	Executable Based Variable
EIF	Executable Infection Flag
EOF	End Offset
EPN	Executable Private Number
ERF	Executable Repair Flag
FAT	File Allocation Table
FBEA	Foreign Block Eradication Algorithm
IAVS	Ideal Anti-Virus System

IBM	International Business Machine
IIB	Image Information Block
IVT	Interrupt Vector Table
LBC	Loader Based Constant
LEP	Load and Execute Program
LOC	Trapdoor mark LOCATION
Max.	Maximum
Min.	Minimum
MIS	Memory Image Size
NAB	Number of Added Bytes
NCB	Number of Changed Bytes
OS	Operating System
PC	Personal Computer
RD	Route Directory
SBD	Simple Block Distribution
SDS	Self-Defense System
SDR	Self-Defense Routine
SIB	Secret Identification Block
SMI	Start of Memory Image
SOF	Start Offset
TGS	Target site/cell Group Size
TMV	Trapdoor Mark Value
TOA	Trapdoor Offset Address
VDC	Virus Deception Capability
VEP	Virus Execution Probability
VFEA	Virus Follower Eradication Algorithm
VGS	Virus Group Size
VIP	Virus Infection Probability

Tab.1

	LEP1		LEP2	
File Name	g	TMV=g+6	g	TMV= g⊕305
HAMED.COM	72+65+77=214	220	65-69= -4	-307
MUNTHER.EXE	77+85+78=240	246	85-84= 1	304
KARIM.SYS	75+65+82=222	228	65-73= -8	-311

Tab.2

Filename	LOC	[LOC]=TMV
HAMED.COM	500	503
MUNTHER.EXE	1024	1027
KARIM.SYS	856	859

Tab.3 Primary SDR Module Standard Format

LOC1+	Size(Byte)	Content	Function/Algorithm Name
00H	4	TMV11	f11(LOC1, LBC11)=g11(LOC1,LBC12)
04H	4	TMV12	f12(LOC1,LBC13)=g12(LOC1,LBC14)
08H	4	CLEN	f13(LEN)
0CH	4	CSN	CHKS-ALG1
10H	LEN	SDR	SDR code/data

Where:

LEN= SDR Module Length In bytes
CLEN= Coded Length
CSN= CheckSum Number
CHKS-ALG= CheckSum Algorithm

Tab.4 Redundant SDR Module Standard Format

LOC2+	Size(byte)	Content	Function/Algorithm Name
00H	4	TMV21	f21(LOC2,LBC21)=g21(LOC2,LBC22)
04H	4	TMV22	f22(LOC2,LBC23)=g22(LOC2,LBC24)
08H	4	CLEN	f23(LEN)
0CH	4	CSN	CHKS-ALG2
10H	LEN	SDR	SDR code/data

Tab.5 Message File Entry Format		
Offset	Size (Byte)	Content
00H	11	Executable Name (E1, E2, ...)
0BH	4	Value of LOC1
0FH	4	Value of LOC2

Tab.6 IIB Standard Format		
Offset	Contents	Description
00H	SMI	Start Address of the executable Memory Image
04H	MIS	Memory Image Size
08H	DIS	Disk Image Size
0CH	EPN	Executable Private Number
0EH	TOA	The SDR Trapdoor Offset Address relative to SMI
12H	LEP-TRAP	Far pointer to the trapdoor of LEP.

Tab.7	
Conditions	Description
(TOA=SOF) AND (MIS-TOA=EOF)	Clean executable
(TOA>SOF) AND (MIS-TOA=EOF)	Virus block before the SDR
(TOA=SOF) AND (MIS-TOA>EOF)	Virus block after the SDR
(TOA>SOF)AND (MIS-TOA>EOF)	CBD Virus

Tab.8 SIB Standard Format		
Address	Size (Byte)	Content
00	2	Length of SIB in bytes
02	?	SIB contents

Tab.9 “SDR Checksum Numbers Table”	
Block Number	CRC-Checksum (CSN)
1	CSN1
2	CSN2
.	.
.	.
N	CSNn

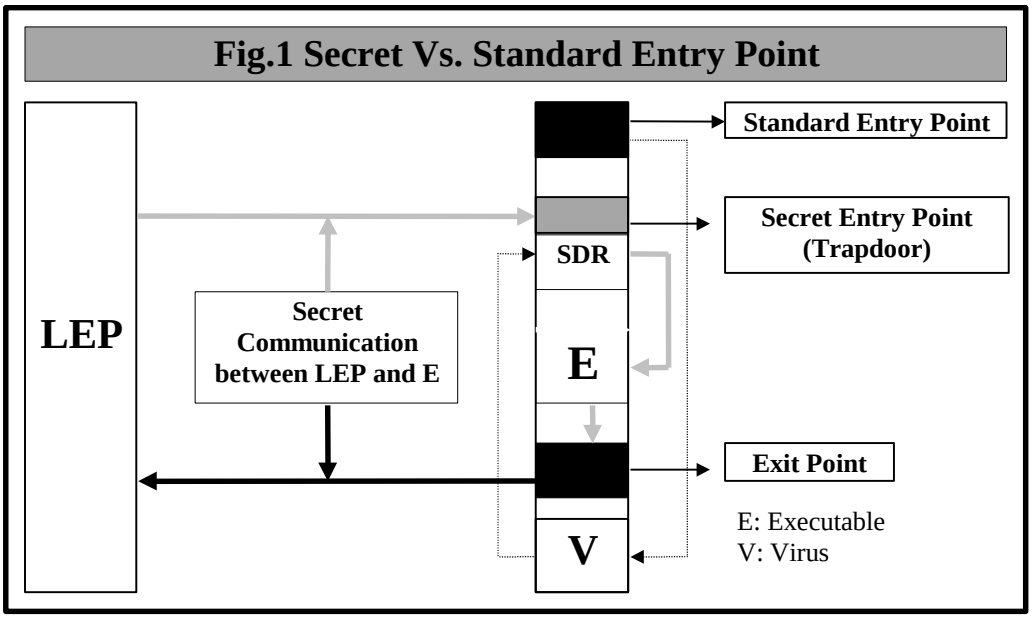


Fig.1 Secret Vs. Standard Entry Point
LEP
Secret Communication between LEP and E
SDR
E
V
Standard Entry Point
Secret Entry Point (Trapdoor)
Exit Point
E: Executable
V: Virus

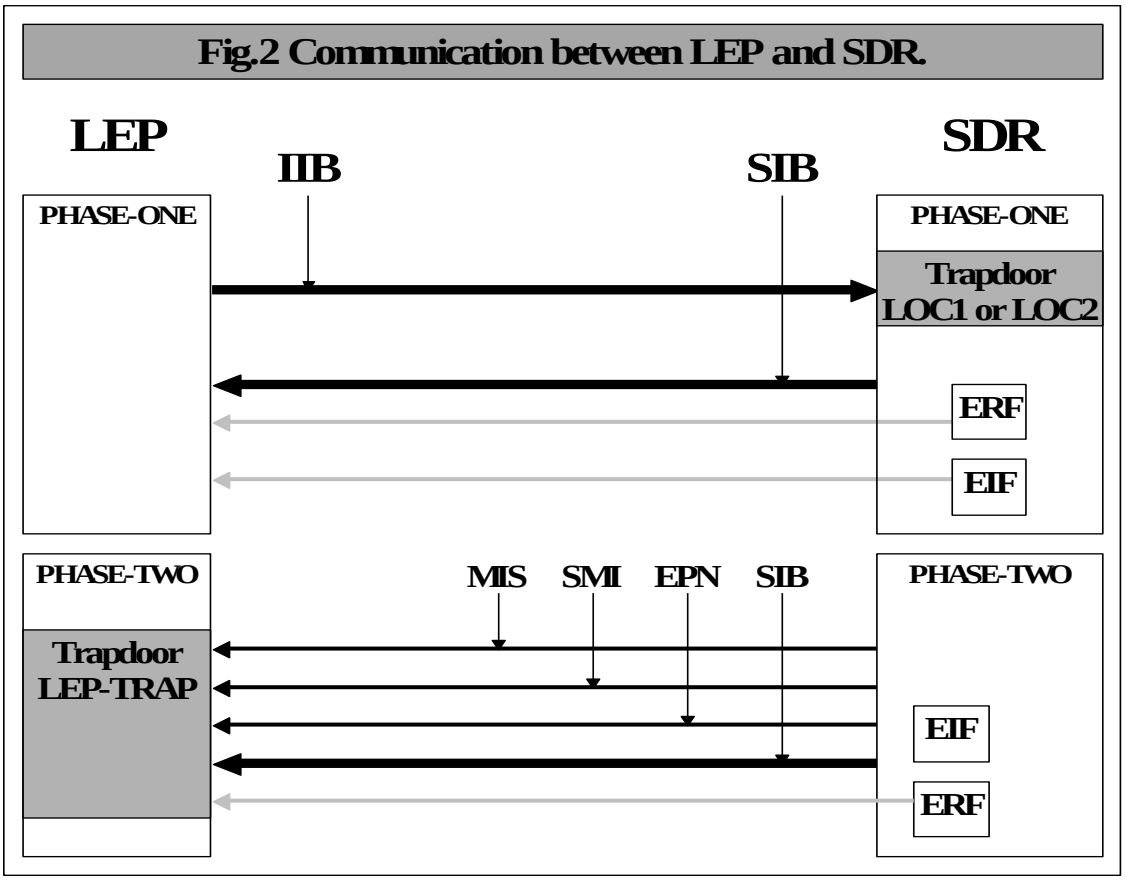


Fig.2 Communication between LEP and SDR.
LEP
SDR
IIB
SIB
PHASE-ONE
PHASE-ONE
Trapdoor LOC1 or LOC2
ERF
EIF
PHASE-TWO
PHASE-TWO
MIS
SMI
EPN
SIB
Trapdoor LEP-TRAP
EIF
ERF

Fig.3 The Protected Executable In Memory

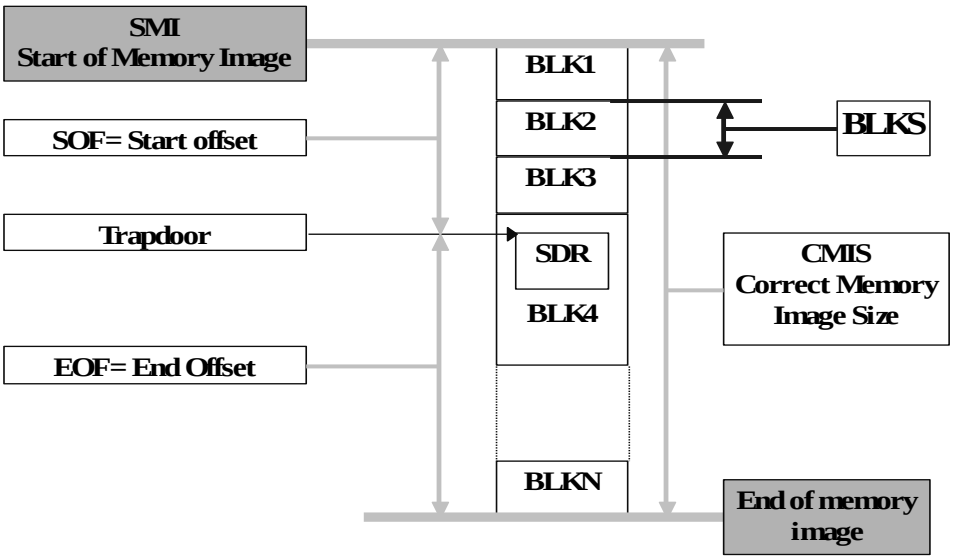


Fig.4 Simple Virus Distribution

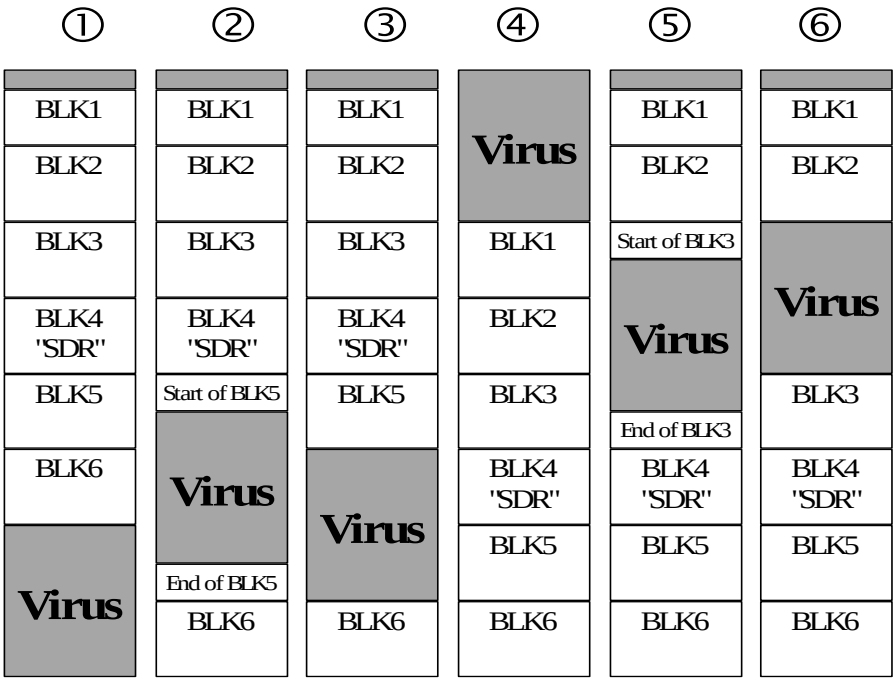


Fig.5 Determining S and E

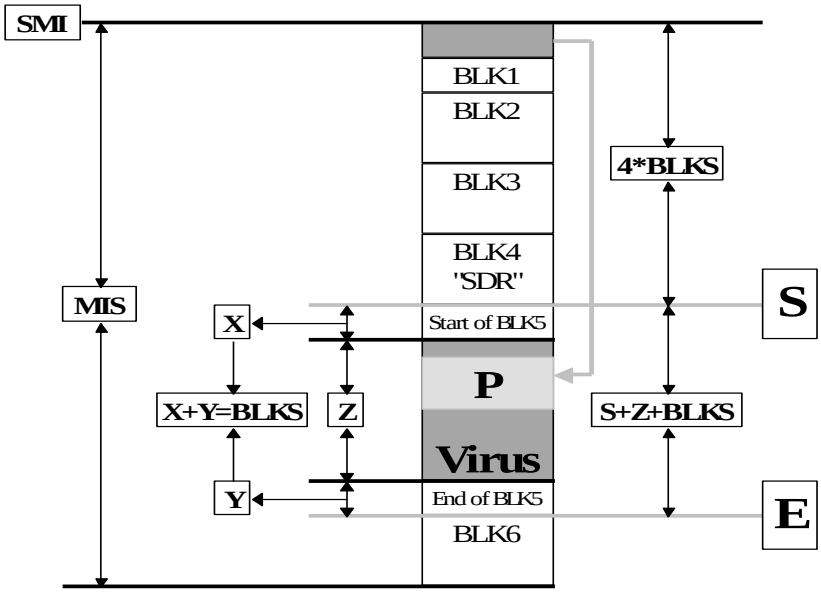
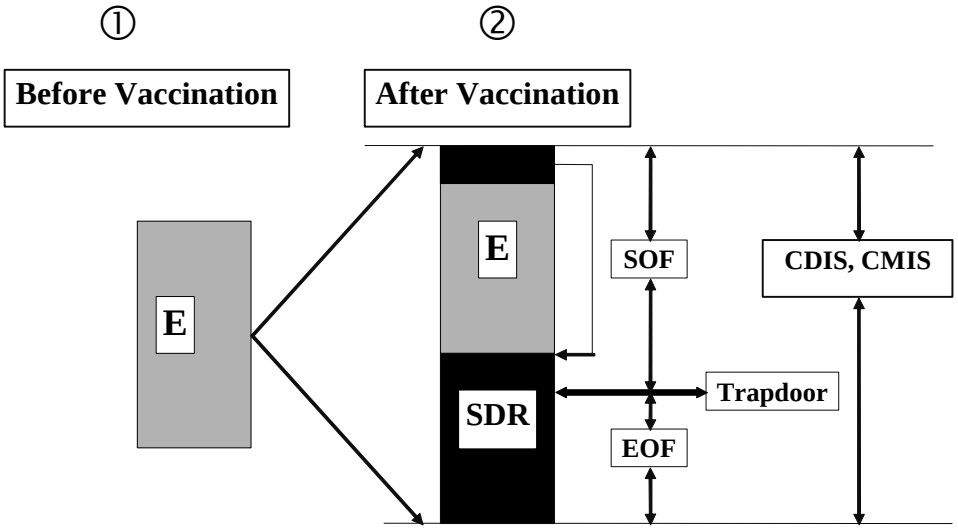


Fig.6 The Executable E Before and After Vaccination.



COMPUTER VIROLOGY: SELF-DEFENSE SYSTEM

Hamid M. A. Abdul-Hussain & Hamed Mizher Shabib
Computer Engineering Dept.
College of Engineering, University of Baghdad

ABSTRACT

Based on the biological human models in defending human body against viruses, a new approach in designing the anti-virus system is introduced. This approach is called SDS(Self-Defence System). The principle of the SDS is that each executable program is responsible of defending itself against viral-attacks. In this system, each executable program is injected with basic anti-virus component which is called Self-Defence Routine. This routine, together with dedicated anti-virus loading program are used to construct the SDS which protects the computer system from virus invasion.

INTRODUCTION

Today, the computer virus pandemic becomes a serious security threat to causal home computers and large corporate networks. Over the years, the anti-virus industry has had to keep pace, as virus writers have become more sophisticated. Therefore, the effort to combat the computer virus must continue until an ideal, universal anti-virus system is designed. Researchers have taken many approaches, and some of the newest and most promising anti-virus technology is modeled on the way the human body fights viruses [1].

Based on the similarities between human and computer viruses (both types of viruses latch onto a host, use its resources to reproduce, and cause a range of symptoms). The objective of this work is to build computer immune system. The proposed system is called Self-Defence System (SDS). The SDS mimics the characteristics of human immune system by distributing the anti-virus components through the computer in the same way the human immune system distributes the anti-bodies. The concept of the SDS is to vaccinate each executable program with special anti-virus component that will detect and eradicate any foreign code attached to the executable. The vaccinated executable is loaded and executed by special centralized anti-virus program which is designed to prevent virus infection and damage to the computer system.

In part this research has been a follow-up on the two papers; the first titled “Computer Virology: Formal Analysis of Computer Viruses” [7], and the second titled “Computer Virology: Toward Designing an Ideal Anti-virus System” [4]. Some terms, classifications, and concepts presented in this research are thoroughly explained in the above papers. For example the efficiency of the SDS is determined according to the analysis criteria described in the second paper mentioned above.

SELF-DEFENCE CONCEPT

The self-defense concept is that: **“Every executable program must be capable of protecting itself against viral attacks by detecting and eradicating any virus attached to its body”.** Accordingly, special part of each executable must be reserved for the self-defence task. This part is called **“Self-Defence Routine (SDR)” and** defined as: **“The part of the**

executable that is specifically designed to protect the executable against viral attacks". SDR is the first basic anti-virus component in SDS, which is embedded, as integral part of an executable E. E will be called the **"protected executable"** along this research. The second basic anti-virus component in SDS is a special OS program called "Load and Execute Program (LEP)". LEP is used to load executables into memory and execute their SDRs secretly by using a special secret communication protocols. LEP and SDR cooperate to satisfy the requirements of the IAVS.

The communication between LEP and SDR consist of two phases (PHASE-ONE and PHASE-TWO). During PHASE-ONE, LEP must transfer control to SDR without giving the attached virus "if any" the opportunity to execute. If SDR detect a virus infection, it will try to eradicate the attached virus and repair the protected executable using the Foreign Block Eradication Algorithm (FBEA). If FBEA failed, SDR must alert LEP to switch to PHASE-TWO. In PHASE-TWO, LEP will execute the virus in a virtual computer system. After the virus execution is completed, SDR will try to eradicate the virus and repair the infected executable using Virus Follower Eradication Algorithm (VFEA). The following sections will discuss how SDR and LEP are designed, and the communication protocols they use in PHASE-ONE and PHASE-TWO.

SECRET ENTRY POINT

Most viruses tacks themselves into an executable program and ensure that they will execute before their host, therefore, they must redirect the executable standard entry point to point to the virus entry point. The standard entry point is the location to which the OS transfers control when deciding to execute the executable. The standard entry point is considered the most vulnerable spot for viral attacks.

In PHASE-ONE, LEP must execute SDR without giving the opportunity to any attached virus "if any" to execute. Clearly, this cannot be done by transferring control to the standard entry point of the executable because viruses know where the standard entry point of the executable is, and they always redirect it to ensure that they will get the opportunity to execute before the infected executable. There is only one way to ensure that SDR will get the opportunity to execute before any attached virus. That is, by using a nonstandard or secret entry point. This kind of entry point is called trapdoor: "**The Trapdoor is a secret entry point to the executable [2]".** Fig.1 shows how the trapdoor can be used.

In fact, the secret entry point will be used to solve the standard entry point problem. Therefore, the trapdoor location must vary from one executable to the other. For example, while the trapdoor of the executable 'E1' is located at offset 'X1', the trapdoor of the executable 'E2' must be located at offset 'X2', where 'X1≠X2'. Because the same standard entry point problem will be faced again, if the trapdoor location is standardized for all executables. For example, if the trapdoor location is standardized to be at offset 'X1' for all executables, then: **"The virus will know that there is another entry point to the executable located at offset 'X1'. Consequently, it will redirect to this entry point to ensure that it will get the opportunity to execute before its host".** Therefore, a new technique that will enable LEP to find or mark the trapdoor while, at the same time, viruses cannot find the trapdoor. The following sections describe how LEP mark the trapdoor, the standard SDR module format, and how different LEPs can exchange protected executables between each other.

TRAPDOOR MARKING

LEP can find the trapdoor location of a given executable by using special type of information stored in the executable itself. This special type of information will be called the "**Trapdoor Mark"** and defined as "**A special code or string of characters whose location (LOC) or value (TMV "Trapdoor Mark Value") can be used to find the trapdoor".**

LEP must calculate the TMV and store it into LOC using criteria known by LEP only. This criterion differs from one LEP to the other. Therefore, viruses cannot estimate in advance which trapdoor mark is used by a randomly selected LEP. For example, if LEP1 and LEP2 are used in two different computers PC1 and PC2 respectively. LEP1 define (TMV= "HMS"), and

define the first byte position which follow the 3'rd character as the trapdoor (Trapdoor=LOC+3), while LEP2 define (TMV=666) and (Trapdoor=LOC+2). Since the virus cannot estimate in advance in which computer it exist, in PC1, PC2, or some other computer, it cannot estimate in advance whether the TMV is, "HMS", 666, or any other number or string. The efficiency and reliability of the secret entry point depend on the TMV type and the method used to generate it. In general, two TMV types can be defined:

FIXED TMV

In this type, the same TMV used for all executables in the system. The generation and searching formulas are very simple in this case. There is, however, a vulnerable spot in this case. Since the TMV is used in all executables, it will be the common code between all executable. Viruses can determine the fixed TMV by comparing two or more executables to see which code is common between them. This code might be the TMV. Note that if each executable in the system uses the suggested self-defence technique efficiently, no virus will get the opportunity to execute and therefore it cannot perform the comparison described above.

Variable TMV

In this type, a different TMV used for each executable in the system. Only LEP knows how to generate and find this TMV for a given executable. Two general techniques can be used to generate a variable TMV:

A- Value-Based Technique: "In this technique, TMV is calculated using a predefined formula and stored in LOC". The general Value-Based generation formula can be defined as follow: **"TMV= f(g(EBV), LBC)".** Where:

- **EBV "Executable-Based Variable":** This value is constant with respect to a given executable. But it differ from one executable to the other. Therefore, LEP can ensure that TMV will change from one executable to the other.
- **LBC "Loader-Based Constant":** This value is constant with respect to a given LEP. But differ from one LEP to the other. This ensures that the TMV will change from one LEP to the other even if both use the same EBV, g, and f.
- **f and g :** Functions (arithmetic or logical relation).

In general, LEP must search for the trapdoor mark location (LOC) using TMV. The general Value-Based searching formula can be defined as follow: **"IF ([P]=TMV) THEN (LOC=P)".** Where:

- **P= Address Pointer**
- **[P]= Content of the location at address P.**

For example, the executable file name can be used as EBV, because it is constant for a given executable file. The TMV generation formula will be: **"TMV= f(g(Executable Name),LBC)".** Let us define the following, for LEP1 and LEP2:

Component	LEP1-Definition	LEP2-Definition
g	ASCII of the 1st Character + ASCII of the 2nd Character + ASCII of the 3rd Character	ASCII of the 2nd character - ASCII of the 4th character
LBC	6	305
f	g + LBC	g ⊕ LBC

Tab.1 shows the TMV generated for three different file names by using the definitions shown above. LEP1 can find LOC for HAMED.COM using the following algorithm:

Calculate :TMV=72+65+77+6=220
P=0
IF ([P]=220) THEN (LOC=P) ; GO TO 6
P=P+1
IF (P<Size of HAMED.COM) THEN 3

End.

B- Condition-Based Technique: "In this technique, TMV is selected so that a predefined condition satisfied". The general Condition-Based generation formula can be defined as follow: **"Select TMV To Satisfy: f(LOC,LBC1)=g(LOC,LBC2)".**

The general Condition-Based searching formula can be defined as follow: **"IF (f(P,LBC1)=g(P,LBC2)) THEN (LOC=P)".** For example, the following can be defined:

LBC1	f(P,LBC1)	LBC2	g(P,LBC2)	Condition
0	[P]+LBC1	0	P+LBC2	[P]=P
1	[P]+LBC1	2	P*LBC2	[P]+1=P*2
19	[P]-LBC1	0	P⊕[P-1]+LBC2	[P]-19=P⊕[P-1]

Tab.2 shows the TMVs generated to satisfy the condition "[P]=P+3" for the three executables shown in Tab.1. The following algorithm can be used by LEP to search for LOC, assuming that f, g, and LBC are defined as shown above:

P=0
IF ([P]=P+3) THEN (LOC=P) ; GO TO 5
P=P+1
IF (P<Size of HAMED.COM) THEN 2
End.

However, it is very important to test the influence of virus infection on the selected condition. For example, if the condition [LOC]=LOC+3 used, and a virus inserts 256-bytes in the start of HAMED.COM and, hence, shift the contents of HAMED.COM by 256-byte. In this case the "TMV=503" will be shifted from location 500 to location 756. Clearly, the condition will be not satisfied in location 500 and location 756. Conditions that link the content of LOC with the contents of its predecessor or successor locations are not affected by the virus infection. For example, the condition:[LOC]=[LOC-1]+[LOC+1].

SDR Module Standard Format

The SDR module consists of the SDR code/data and trapdoor marking information. Two problems must be considered before defining the standard format of the SDR module:

Pseudo Trapdoor Mark: "The pseudo trapdoor mark is any value satisfy the searching formula and cause LEP to transfer control to an incorrect location". The increased probability of pseudo trapdoor mark in the value-based technique is the great disadvantage of this technique. For example, in Tab.1 the TMV associated with HAMED.COM is 220 with respect to LEP1. Clearly, this value may exist more than one time within the code or data of HAMED.COM. Moreover, viruses may consider this as a vulnerable spot for attack. For example, TMV will range from 0 to 255 if represented as one byte storage location. The virus can insert a table that contains all of the expected values (from 0 to 255) in the start of the infected executable. When LEP search the infected executable, it will find a pseudo trapdoor mark in the virus table and transfer control accordingly. This transfer will activate the virus to start its execution.

The probability of finding a pseudo trapdoor mark in the condition-based technique is less relative to the value-based technique. Because it searches for a **value that satisfy a condition.** For example, in Tab.2, even though the value 503 may exist many times within the code or data of HAMED.COM, it will be considered as a trapdoor mark only if it exists at offset address 500.

In general, the solution to the pseudo trapdoor mark problem is to use two TMVs associated with special integrity check information. Tab.3 shows the suggested standard format of the SDR module. According to Tab.3, LEP can find the SDR module as follows:

1. **P=0**

2. **IF (f11(P,LBC11)=g11(P,LBC12)) THEN 6**
3. **P=P+1**
4. **IF P<Executable_Size THEN go to 2**

;Trapdoor not found,

5. **Display Alarm Message "Cannot Find the Trapdoor".**

;Test if the second condition satisfied

6. **IF (f12(P,LBC13)=g12(P,LBC13)) THEN go to 8**

;TMV2 Not exist, continue the search

7. **Go To 3**

;Perform checksum test

8. **LEN= $f13^{-1}$([P+08H])**
9. **Using CHKS-ALG1, Set C= Checksum of the block (P+10H+LEN)**
10. **IF C=[P+0CH] THEN go to 12**

;Checksum Error, Continue The search

11. **Go To 3**

;Trapdoor Mark Found

12. **LOC1=P**

Destruction: Destruction can be caused by a virus infection. If the virus, for example, inserts part or all of its code between TMV11 and TMV12, between the header and the SDR module, or inside the SDR module. The solution to this problem is to use redundancy. That is, two SDR modules must be used so that if one of them destroyed the other one can be used. The first SDR module (shown in Tab.3) will be called the **"Primary SDR Module"** and the second SDR module will be called the **"Redundant SDR Module"**. The redundant SDR module is shown in Tab.4.

LEP-To-LEP Communication

As mentioned earlier, each LEP uses its own standard to generate and search for TMV and LOC. Users may ask what happens when the protected executable 'E' copied from the environment of LEP1 to the environment of LEP2. That is: **"How LEP2 can find LOC1 and LOC2 of 'E', without giving viruses the opportunity to find them".** Clearly, there must be a standard and secure communication protocol between LEP1 and LEP2. There are many ways to do this.

First, the communication can be done by using a global TMV, say "HMS". When the user asks LEP1 to copy 'E' to be used by LEP2, LEP1 must store "HMS" at LOC1 and LOC2. When the user asks LEP2 to execute 'E' for the first time, it will search 'E' for the global TMV (i.e. HMS) to find LOC1 and LOC2. Once "HMS" is found LEP2 will recalculate LOC1, LOC2 and reorganize the SDR locations according to its formula. The advantage of this method is its transparency for users. The great disadvantage of this method is its vulnerability. Viruses know that the global TMV used in copies is "HMS", therefore, they can easily find LOC1 and LOC2 of the protected executable.

Second, the communication can be done through the computer user. The user can ask LEP1 about the value of LOC1 and LOC2 of 'E' after copying the executable to a new diskette. The user must give the value of LOC1 and LOC2 to LEP2 (i.e. manually) when he wants to execute 'E' for the first time. The advantage of this method is that it is secure, because there is no way that a virus will know the value of LOC1 and LOC2. The disadvantage of this method is that it depends on the computer user responsibility. A serious problem arises when the user decides to copy a large number of executables. For example, if the user wants to copy 100-executable from DISK1 to DISK2, then, he must ask LEP1 about the value of LOC1 and LOC2 for each one of these executable, keep these (200-Value) values in his mind, and give them to LEP2.

Third, the best suitable method is to use Message-Files. When the user asks LEP1 to copy the 100-executable (E1, E2,… E100) from DISK1 to DISK2, LEP1 will ask the user about

the message file name to store the values of LOC1 and LOC2 in it. Assuming that the user uses the name MESSF.LOC. LEP will create a message file with the name MESSF.LOC in DISK2, and insert 100-entry in this file. Each entry represent one executable. The standard entry format is shown in Tab.5. LEP2 can use the message file to find LOC1 and LOC2 of each executable individually or all executables at once. Regardless of the number of executables being copied, the user needs to remember only the message file name.

-The Communication Standards

Before describing the communication protocols used in PHASE-ONE and PHASE-TWO; and the operations of LEP and SDR in each phase, the standard data blocks, variables, and flags used in the system must be defined and described. Fig.2 shows an overview of the communication system and its individual components. Each data block, variable, and flag used for specific purpose. The following sections will describe these components and explain the purpose of using them.

- Image Information Block (IIB)

In PHASE-ONE, LEP must prepare the Image Information Block (IIB) before transferring control to the SDR. The IIB describes the status of the disk and memory images of E. The standard format of the IIB is shown in Tab.6.

MIS=DIS for COM and SYS files. But MIS≠DIS for EXE and OVL files, because the header exist in the disk image but not loaded with the memory image. MIS and DIS are used by the detection algorithm. TOA is used by the detection and eradication algorithms as will be explained later. EPN (Executable Private Number) is generated by LEP for each executable before executing it. EPN of E is considered, by LEP, as the “identifier” or “Secret Name” of E. While in PHASE-TWO, the SDR of E must pass the EPN of E associated with the other identification information to LEP, so that it can get the permission from LEP to access the disk image of E. LEP-TRAP is a pointer to a special trapdoor within LEP. This trapdoor is used to ensure that the SDR can communicate with LEP secretly while in PHASE-TWO.

- Detection and Eradication Information

The SDR must know the following information about the protected executable:

1- **Executable Critical Bytes: “The critical byte is any byte which is virtually guaranteed to be changed after a virus infection”.** The executable size, the first three bytes of a COM or SYS, and the EXE or OVL file header are considered as critical bytes. Determining the critical bytes of a given executable, require a deep understanding of how the computer virus infect it. In general, the following components are calculated and stored in the SDR of E during the protection process “see section 5”:
   1. **CMIS= Correct Memory Image Size (CMIS)**
   2. **CDIS= Correct Disk Image Size (CDIS)**
   3. **The correct values of any other critical byte. Such as the first three bytes of COM and SYS files and the header of EXE and OVL files.**

- **Integrity Check Information (Checksum):** The SDR is assumed to view the executable memory image as group of N-Blocks (BLK1,BLK2, ... BLKN) “see Fig.3”. The block size (BLKS) is equal for all blocks and stored in (BLKS). A CheckSum Number (CSN) calculated for each block using the algorithm CHKS-ALG, and stored in a special SDR table. Also, a given block ‘BLKi” is assumed to be valid (i.e. has a correct checksum) if the following condition satisfied: **“CHKS-ALG(BLKi, CSNi)=0”.**
- **Position Test Information:** The position test information (SOF and EOF) are used (with TOA, and MIS) by the FBEA to find the virus block position relative to the SDR module as show in Tab.7. Fig.3 shows how the protected executable appear in memory. SOF and EOF represent the position test information and defined as follow:

- **The Start Offset (SOF):** Is the offset of the SDR trapdoor relative to the start of the memory image**.**
- **The End Offset (EOF):** Is the backward offset of the SDR trapdoor relative to the end of memory image**.**

- SDR Internal Flags

SDR uses two internal flags:

1- Executable Infection Flag (EIF): SDR set this flag if it detects a virus infection. The status of this flag is returned to LEP at the end of PHASE-ONE.

2- Executable Repair Flag (ERF): SDR set this flag if it can repair the infected executable. The status of this flag is returned to LEP at the end of PHASE-ONE, and the end of PHASE-TWO through LEP-TRAP.

2.4 The Secret Identification Block (SIB)

The SIB passed by the SDR of E to LEP during PHASE-ONE. When LEP switch to PHASE-TWO, it can use SIB to distinguish the SDR of E from the SDRs of the other executables and viruses. The standard format of SIB is described in Tab.8.

3- PHASE-ONE Communication

PHASE-ONE starts when LEP receives a request to execute an executable. Assuming that LEP receives a request to execute the executable E, the following sections describe the sequence of operations:

Locating and Executing SDR

The following steps describe loading the executable by LEP and preparing the IIB:

Step-1: "Loading the executable memory image"

1. **Find the disk image of E.**
2. **Store DIS in the IIB.**
3. **Assign EPN to E and store it in the IIB.**
4. **Store LEP-TRAP in the IIB**
5. **Reserve a memory block to store the executable memory image. And store the start address (SMI) in the IIB.**
6. **Load the executable memory image into the reserved block.**
7. **Store the MIS in the IIB.**

Step-2: "Finding the Trapdoor"

1. **Search for the trapdoor of the primary SDR module.**
2. **IF (the trapdoor found) THEN go to 6**
3. **Search for the trapdoor of the redundant SDR module.**
4. **IF (the trapdoor found) THEN go to 6**
5. **LEP cannot find any one of the trapdoors. This can happen if the executable is not Self-Protected, the executable is new in the system, or both SDR modules are destroyed due to a virus infection. In either case, LEP must display alarm message to the user, and ask him what to do. Depending on the user response, LEP must proceed as follow:**
   - **If E is not self-protected, then execute it in a virtual computer (explained later).**
   - **If E is new, then prepare it using the message file.**
   - **If E destroyed by a virus infection, then, avoid executing it**
6. **Store the offset of the trapdoor in TOA.**
7. **Store TOA in the IIB.**

Step-3:"Execute and Wait"

LEP can transfer control to the SDR trapdoor using a far call instruction, and wait until the SDR return control again. What happens when the SDR receives control is explained in the next section.

SDR Operation in PHASE-ONE

The SDR operation can be described by two cycles: **"Self-Test Cycle, and Self-Repair Cycle"**. Both cycles are described in the following sections.

Self-Test Cycle

This cycle initiated each time SDR executed, in this cycle SDR must decide whether E is infected or not. If E is infected, SDR will set EIF, and clear it otherwise. The status of EIF returned from SDR to LEP as shown in Fig.2. The following algorithm describe the self-test cycle:

;Size test
IF DIS≠CDIS THEN go to 13
1. IF MIS≠CMIS THEN go to 13
;Critical byte test
2. IF (Any critical byte changed) THEN go to 13
;Checksum test
3. i=1
4. C=CHKS-ALG(BLKi, CSNi)
5. IF C≠0 THEN go to 13
6. i=i+1
7. IF i ≤ N THEN go to 5
;Position test
8. IF SOF≠TOA THEN go to 13
9. IF EOF≠MIS-TOA go to 13
"The executable is clean".
10.EIF=0
11.Return to LEP.
"The executable is infected".
12.EIF=1
13.Go to the Self-Repair Cycle.

Self-Repair Cycle

In this cycle, SDR will try to eradicate the virus and repair the protected executable. SDR will set ERF if the protected executable repaired properly and clear it otherwise. A new eradication algorithm (Foreign Block Eradication Algorithm "FBEA") will be used. FBEA capability depends on how the virus distributes itself within the infected executable. The FBEA described below can eradicate SBD viruses efficiently, and can be upgraded to eradicate CBD viruses as well. However, because the CBD idea not used by viruses yet "see [3]", the discussion will be limited for the SBD viruses only. Assuming a simple virus distribution, the eradication algorithm can be divided into the following steps:

Step-1: "Find the virus block position relative to the SDR"

In general, if the virus block inserted after SDR, the infected executable memory image will take the form of image ①,②, or ③ in Fig.4. Therefore, the search must start from BLK1. If the virus block inserted before SDR, the infected executable memory image will take the form of image ④,⑤, or ⑥. Therefore, the search must start from BLKN. Finding the virus block position relative to SDR can be done by using the position test information TOA, MIS, SOF, and EOF as shown in Tab.7.

Step-2: "Starting the block checksum test"

For example, let us assume that the virus is inserted after the SDR, therefore, the block checksum test must start from BLK1. The following algorithm can be used in this case:

1. **i=1**
2. **IF BLKi contain any critical byte, then repair the critical bytes.**
   "Only BLK1 in image ①, ②, and ③ affected by this step"
3. **C=CHKS-ALG(BLKi, CSNi) ;Calculate the checksum of BLKi**
4. **IF C≠0 THEN go to Step-3 ;Invalid block**
5. **Move BLKi into BUF ;Valid block**
6. **i=i+1**
7. **IF i ≤ N THEN go to 2**

Arriving to this point means that the virus has inserted all of its added bytes at the end of the infected executable, as shown in image ①.

Go to Step-6

"Note the difference between (go to 3) and (go to Step-3)"

Step-3: "Reversing the block checksum test order".

Arriving to this step means that the virus inserts its added bytes after SDR but not at the end of the infected executable. In this case, the infected executable memory image expected to take the form of image ② or ③. Because BLKi is not found, the searching processes must be reversed. The following algorithm can be used:

1. **j=0**
2. **m=N-j**
3. **IF BLKm contain any critical byte, then, repair the critical bytes.**
4. **C=CHKS-ALG(BLKm, CSNm) ;Calculate checksum number of BLKm**
5. **IF C≠0 THEN go to Step-4 ;Invalid checksum**
6. **Move BLKm into BUF ;Valid checksum**
7. **IF m=i THEN go to 10**
8. **j=j+1**
9. **Go to 2**

Arriving to this point means that the virus block inserted between two consecutive blocks (BLKi-1 and BLKi) without destroying any one of them, as shown in image ③. Because the blocks (BLK1, BLK2, ..., BLKi-1) are moved into BUF in Step-2. And the blocks (BLKi, BLKi+1, ... BLKN) are moved into BUF in Step-3. Therefore, all blocks are moved into BUF.

10. **Go to Step-6**

Step-4: "Repairing the damaged block"

Arriving to this point means that the virus has inserted its block inside BLKi and, hence, destroyed this block. This is shown in image ②. Two routines to repair the destroyed block BLKi will be discussed:

REPAIR1:

Inputs:

- **i= Block number**
- **S= Start offset address of BLKi which contain the virus block.**
- **E= End offset address of BLKi which contain the virus block.**
- **Z= MIS-CMIS=NAB**
- **P= Offset address of a byte that is guaranteed to exist in the virus block.**

Outputs:

- **C=0 ;BLKi cannot be repaired**
- **C=1 ;BLKi repaired properly and stored in the buffer BBK**

The Idea: REPAIR1 assumes that the virus block start at P and define: **"VS (Virus-Start)=P, and VE (Virus-End)= VS+Z-1".** And then, moves the bytes at block (S-To-(VS-1)) and block ((VE+1)-To-E) into the buffer BBK. If the checksum test on BBK fails, REPAIR1 assumes that P is not the actual start of the virus block. That is, there is at least one byte belong to the virus block and exist before P. Therefore, REPAIR1 shift VS and VE up by one byte position, and repeat the test process.

Algorithm:

1. **VS=P**
2. **VE=VS+Z-1**
3. **IF {VE>(E-1)} THEN {[VS=VS-(VE-(E-1))] AND [VE=VS+Z-1]}**

   "This ensures that VE≤(E-1). Note that, the byte at E must belong to BLKi, because, otherwise, all of the content of BLKi exists above the virus block. And this can happen only if BLKi wasn't destroyed by the virus infection".
4. **Move the bytes at block (S-To-(VS-1)) into BBK**
5. **Move the bytes at block ((VE+1)-To-E) into BBK**
6. **C=CHKS-ALG(BBK, CSNi)**
7. **IF C= 0 THEN go to 12 ;Valid block**

   "Arriving to this point means that VS is not the actual start of the virus. That is, there is at least one byte belong to the virus but exist above VS. Therefore, VS must be decremented and the checksum must be calculated again".
8. **VS=VS-1**
9. **VE=VS+Z-1**
10. **IF VS>S THEN go to 4**
11. **C=0, and RETURN "BLKi cannot be repaired"**
12. **C=1, and RETURN "BLKi repaired properly and stored in BBK"**

REPAIR2:

Inputs:

- **i= Block number**
- **S= Start offset address of BLKi which contain the virus block.**
- **E= End offset address of BLKi which contain the virus block.**
- **Z= MIS-CMIS=NAB**

Outputs:

- **C=0 ;BLKi cannot be repaired**
- **C=1 ;BLKi repaired properly and stored in the buffer BBK**

The Idea: REPAIR2 is based on the fact: **"If BLKi *destroyed* by one sequential virus block, then, the byte at S and the byte and E must belong to BLKi. Because otherwise, if the byte at S(E) belong to the virus block, the entire virus block must exist above (below) BLKi, and this can happen only if BLKi wasn't *destroyed* by the virus block insertion".** REPAIR2 moves one byte starting from S and (BLKS-1) bytes starting from E into BBK. If the checksum test on BBK fail, REPAIR2 repeat the process by moving 2-bytes starting from S and (BLKS-2) bytes starting from E into BBK.

Algorithm:

1. **k=S**
2. **j=k+Z+1**
3. **Move the bytes at block (S-To-k) into BBK.**
4. **Move the bytes at block (j-To-E) into BBK.**
5. **C=CHKS-ALG(BBK, CSNi)**
6. **IF C=0 THEN go to 10;Valid block**
7. **k=k+1**
8. **IF k< (BLKS-1) THEN go to 2**

9. **C=0, and RETURN "BLKi cannot be repaired"**
10. **C=1, and RETURN "BLKi repaired properly and stored in BBK"**

The advantage of REPAIR2 over REPAIR1 is that it can work without using P. However, the number of trails or the time needed by REPAIR1 is, in general, less than the time needed by REPAIR2. Which routine "REPAIR1 or REPAIR2" the SDR must use, depend on the prepared input arguments. The input arguments can be prepared as follow:

1. **S= (i-1)*(BLKS) ;The number of bytes in all blocks before BLKi.**
2. **E= S+Z+BLKS;To understand how S and E calculated see Fig.5.**
3. **The SDR can use REPAIR1 only if it can prepare P. Otherwise, it must use REPAIR2. Two ways are suggested to determine P:**
   a) **The entry point of the virus code within BLKi can be found from the standard entry point of the executable. For example, for COM files the displacement of the near jump instruction which is found in the first three bytes of BLK1 can be used to find the virus entry point. After finding this entry point, set P= Virus code entry point.**
   b) **If the virus code entry point cannot be found, the following fact can be used: "IF (Z>BLKS) THEN (The point at "S+(E-S)/2" must be in the virus block)". Therefore: IF (Z>BLKS) THEN (P=S+(E-S)/2)".**

If P prepared, the SDR can proceed as follows:

1. **CALL REPAIR1**
2. **If C=1 THEN go to 6 ;BLKi repaired properly using REPAIR1**
3. **CALL REPAIR2**
4. **IF C=1 THEN go to 6 ;BLKi repaired properly using REPAIR2**

BLKi cannot be repaired using REPAIR1 or REPAIR2.

5. **Go to Step-5**
6. **Move the contents of BBK into the gap of BLKi in BUF.**
7. **Go to Step-6**

Step-5: "Return Error Code and SIB to LEP".

Arriving to this step means that the SDR cannot repair E properly. Therefore, the SDR must return the following information:

1. **EIF=1 ;The file is infected**
2. **ERF=0 ;The file is not repaired**
3. **SIB**
4. **Go to Step-7.**

Step-6: "Repair the Infected Disk Image of E"

Arriving to this step means that BUF contain all blocks (BLK1, ... BLKN) of the protected executable. The content of BUF represent a clean memory image of the executable, therefore, it can be used to repair the infected disk image of E. The following algorithm describes how this can be done.

1. **If the infected executable is an EXE or OVL file, then, put the correct file header in the start of BUF.**
2. **Overwrite the disk image of E by the content of BUF.**
3. **ERF=1 ; Executable Repaired Properly.**
4. **Go to Step-7**

Step-7: "Return To LEP"

The SDR can return control to LEP now. PHASE-ONE terminated at this point.

- PHASE-TWO Communication

LEP starts PHASE-TWO after the SDR execution, in PHASE-ONE, is completed. Once in PHASE-TWO, LEP will test the status of EIF and ERF returned from SDR. The executable is clean if (EIF=0). Therefore, LEP can execute the executable through its standard entry point. The executable was infected but repaired properly by SDR if (EIF=1 and ERF=1). Therefore, LEP must load the executable disk image again. And, then execute it through its standard entry point.

The executable is infected and SDR cannot repair it if (EIF=1 and ERF=0). Therefore, the Virus Follower Eradication Algorithm (VFEA) must be used to eradicate the virus. VFEA is based on the fact that shell type viruses, "See [7]", always repair the infected executable memory image before executing it. Therefore, VFEA use the following eradication approach: **"The infected executable must be executed through its standard entry point so that the attached virus will get the opportunity to execute. The virus will repair the memory image of the executable and execute it. SDR will receive control again and test the memory image. If the memory image repaired properly by the virus, SDR must alert LEP to use the current memory image of the executable to replace the infected disk image of the executable".**

However, giving the opportunity to the virus to execute is very critical. Therefore, before executing the virus, LEP must prepare a trusted environments "Virtual Computer System" to ensure that the virus cannot cause any damage or infection. Preparing the virtual computer system and executing the SDR in PHASE-TWO is described in the following sections:

The Virtual Computer System

The idea of the virtual computer system is described in [4]. In general, the virtual computer system must be designed to satisfy the following requirements:

1- **Protecting System Disks:** Prevent the virus from infecting or destroying any target site/cell in the system disks.
2- **Protecting System Memory:** Prevent the virus from reserving memory space and hide itself there, that is, stay resident in system memory.
3- **Deceiving the Virus:** The virtual computer system must give the virus the illusion that it is running in a normal system. Because, if the virus knows that it exists in a virtual computer system, it may try to use special methods to bypass or deceive the virtual computer system; or it may terminate its execution without repairing its host, therefore, VFEA cannot repair the protected executable.

Satisfying these requirements depend on the computer system hardware (i.e. Real-Mode or Protected-Mode PC) and software (i.e. DOS or WINDOWS). As a case study, preparing a virtual computer system in DOS machines will be explained in what follows:

1- **Protecting System Disks:** In order to prevent the virus from infecting/destroying target sites/cells in the system disks, it must be prevented from writing to these disks. First of all, LEP must ensure that all of the disk related interrupt vectors points to special handlers. Therefore, LEP must redirect the following vectors:
   1. **Vector 13H (BIOS Disk Interrupt INT 13H).**
   2. **Vector 21H (DOS-API Interrupt INT 21H).**
   3. **Vector 25H (DOS: Absolute Disk Read Interrupt INT 25H) "see [5]".**
   4. **Vector 26H (DOS: Absolute Disk Write Interrupt INT 26H).**

   Under DOS, if the virus knows that it exists in a virtual computer system, it may try to use direct hardware access method to bypass the virtual computer handlers. The virtual computer can use one of the following methods to prevent viruses from accessing the disk: First, the handler of the virtual computer can reject any write to disk request by returning some error code which indicates that the requested operation

cannot be performed because, for example, the drive is not ready. Rejecting all writes to disk requests can help the virus in deciding whether it is working in a virtual computer system or not. Second, the virtual computer system can trick the virus to believe that the requested write to disk operation is performed while it is not. This can be done by returning a no error code that indicates the requested operation performed properly, without performing the actual operation. If this method used, the virus can decide whether it is working in a virtual computer system or not by using the following trick:

1. **Send a request to write the data block (BLK) to disk (C:).**
2. **Read the data block from disk (C:) into (BLK1)**
3. **If BLK≠BLK1, then, the data wasn't written to the disk. Therefore, the system is a virtual computer system.**

Finally, all disk read/write operation can be redirected to a special RAM disk instead of the actual disk. In this way, the virtual computer system can trick the virus to believe that the requested write to disk operation was performed and the data written to the disk properly, while the data was written to the RAM disk. If the virus request the data later, the virtual computer can read it from the RAM disk. This means that the virtual computer must handle both disk read and write operations, this explains why vector 25H was redirect above.

2- Protecting the IVT: The virtual computer system must save the content of the IVT before executing the virus, and restore it once the virus execution completed. This ensures that any redirection to the IVT vectors by a resident type virus is fixed once the virus execution is completed.

3- Protecting System RAM: Protecting the PC system RAM can be done as follow:

- **Save the amount of the available URAM (the word at 0:0413H) before executing the virus. And restore the content of 0:0413H once the virus execution completed. This ensures that the virus cannot decrement the amount of URAM and install itself at the end of the URAM.**
- **If the virus tries to allocate a memory block using function 48H of INT 21H, then, the virtual computer must store the address of the allocated memory block so that it can release this block once the virus execution completed. Therefore, it must redirect vector 21H.**
- **Save a map of all memory control blocks "see "[٥] before executing the virus. And restore them once the virus execution completed. This ensures that the virus cannot allocate memory by directly accessing the memory control blocks.**
- **Redirect vector 66H to special handler before executing the virus. INT 66H is used to access the Expanded Memory Manger functions. This interrupt must be handled only if the system uses expanded memory. If the virus allocate a page (or pages) in the expanded memory, the handler must store the handles reserved by the expanded memory system for the allocated pages. This handle can be used to release the reserved pages when the virus execution completed. Expanded memory expands RAM beyond the 640KB limit, for more information see [5].**
- **Extended memory exists in AT machines (the memory beyond the 1 MB limit "see [5]"). Because the eXtended Memory Manager functions are called through a FAR CALL instruction, instead of the special interrupt, the virtual computer cannot intercept requests to the extended memory function. Therefore, the virtual computer must ensure that all of the extended memory is free before executing the virus. And free it again after the virus execution completed, to ensure that viruses cannot allocate memory in the extended memory area.**

Execute SDR and Wait

After preparing the virtual computer system, LEP can execute E through its standard entry point and wait until it receives a special call which indicate that the virus execution was completed and SDR was executed. Note that LEP can switch from the virtual computer system back to the normal system only after ensuring that the virus execution was completed. Receiving the call is the clue that indicates to LEP that SDR is the currently active program. Because viruses always try to disable or circumvent the protection mechanism, the following requirements must be satisfied:

1- **The virus cannot mask the call:** Therefore, in the PC system the software interrupt mechanism cannot be used to perform the call. Because the virus can redirect all of the interrupt vectors to its own handler, therefore, it can mask the call and preventing it from arriving to LEP. The call must be direct, that is, by using jump or call instructions. This explains why LEP passes a pointer (LEP-TRAP) to its trapdoor to SDR in PHASE-ONE.

2- **The virus cannot deceive LEP:** In order to deceive LEP by a tricky call, the virus must know where to send the call and which information to pass with it. Therefore, the virus cannot deceive LEP through LEP-TRAP, because it doesn't know the correct values of LEP-TRAP, EPN, and SIB. Since these variables passed from LEP to SDR in PHASE-ONE and the virus was inactive at that time, therefore, there is no way to know these variables by the virus.

SDR Operation in PHASE-TWO

When the SDR receives control in PHASE-TWO it will proceed as follows:

1. **IF EIF=1 THEN go to 3**
   ;E is clean
2. **Continue the execution of E normally**
   ;E is infected, start the repair cycle using VFEA
   ;Perform checksum test
3. **i=1**
4. **C=CHKS-ALG(BLKi, CSNi)**
5. **IF C≠0 THEN go to 12**
6. **i=i+1**
7. **IF i≤N THEN go to 4**
   ;All blocks are valid.
8. **ERF=1 ;The memory image is repaired properly**
9. **SMI= Start of the repaired Memory Image**
10. **MIS= Size of the repaired Memory Image**
11. **Go to 15**
    ;BLKi is invalid
12. **ERF=0 ;The memory image cannot be repaired**
13. **Prepare SIB**
14. **EPN= Executable Private Number that was received from LEP in PHASE-ONE**
    ;Transfer control to LEP through LEP-TRAP
15. **JMP FAR PTR LEP-TRAP**

Receiving Control from SDR

The virtual computer system prevents any program from writing to the system disk to ensure that viruses cannot replicate themselves. However, the question is how the SDR of E can repair the infected disk image of E. LEP can give the permission to the SDR to perform disk write operations. SDR must call LEP through LEP-TRAP with the proper EPN and SIB to get this permission. Giving this permission to SDR can be done in one of two ways:

1- **Filtering:** In this case, the virtual computer must be capable of distinguishing between the SDR write to disk requests and the other requests. This can be done, for example, by sending EPN and SIB with each request.

2- **Restoring the IVT:** LEP can restore the IVT content when receiving EPN and SIB through its trapdoor. Therefore, SDR can access the disk as desired by using DOS and BIOS services.

However, giving the permission to SDR to access the disk is not a good idea, because, the virus exist in system memory and may, in some way, interfere with the SDR operation, so there is a possibility that the virus will get the opportunity to access the disk. Therefore, the following method is suggested: **"Instead of giving the disk access permission to SDR, LEP can repair the disk image by itself after receiving the necessary information from SDR".**

Therefore, as shown above, SDR return SMI and MIS to LEP through LEP-TRAP. After receiving this information, LEP can proceed as follows:

1. **Compare the received EPN with the one that was given to SDR in PHASE-ONE, if not equal go to 6**
2. **Compare the received SIB with the one that was received from SDR at the end of PHASE-ONE, if not equal go to 6**
3. **IF ERF=0, THEN, switch to normal mode "SDR cannot prepare a clean backup image".**

   "A clean backup image is available and LEP must use it to replace the infected disk image, as follows:"
4. **Use the MIS-byte block that starts at (SMI) to replace the executable whose private number is EPN.**
5. **Switch to normal mode.**
6. **The SDR identification information (EPN or SIB) are invalid, therefore, reject the call, and display alarm message.**

4.5 Switching to Normal Mode

LEP can switch to normal mode after receiving the SDR request through LEP-TRAP and repairing the infected disk image if necessary. LEP must do the following so that it can switch to normal mode:

1. **Restore the IVT**
2. **Release any memory block allocated during the virus execution using DOS function 48H, URAM, memory control blocks, expanded memory, and extended memory.**
3. **Release the memory allocated for E, to ensure that the viral code which exist at this memory area will be destroyed.**
4. **Release the RAM disk memory area.**

5- Vaccination

There is a large number of executables in the world. All of these executables are developed without any built-in SDR. Clearly, some way must be found to convert these executables to self-protected executable. Vaccination can be used to do this. In this research, vaccination is defined as: **"The process of converting an executable into a self-protected executable by injecting the SDR module inside it".**

"**Note:** Some references (see [6]) refer to 'Vaccination' as one of the protection methods that was suggested by IBM. Some anti-virus programs are designed to inject themselves inside the executable files, and then operate like a '**Benign Virus**' when the vaccinated program executed. The injected anti-virus creates a signature (finger-print) of uninfected executable, and display an alarm if the signature or executable size changed. Unfortunately, the alarm generated after the virus execution. The difference between the suggested SDR and the Vaccine of IBM, is that the SDR executed before any attached virus and uses sophisticated techniques to detect and eradicate the virus".

In fact, the idea of vaccination is borrowed from viruses. A special program "called the INJECTOR" will inject the executable by the vaccine (i.e. SDR). However, there is no searching, or infection routine in this vaccine, therefore, the vaccine cannot replicate itself. The word "injection/inject" is used instead of the word "infection/infect" to distinguish vaccination from virus infection. The following steps can describe the algorithm used by INJECTOR to inject SDR into an executable E.

Step-1: "Injecting SDR in the executable E"

This is similar to what the virus infection routines do, that is:

1. **Store the contents of the standard entry point of E in SDR.**
2. **Attach SDR to E.**
3. **Redirect the standard entry point of E to point to the entry point of SDR.**

After this injection, E will take the form of image ② in Fig.6. From now, image ② is considered the correct executable image. If SDR receives control through the standard entry point and after the completion of its execution, it can repair the image of E to take the form of image ① so that it can execute properly.

Step-2: "Preparing the SDR Module"

As shown in Fig.6, the correct image of E is image ② from the SDR point of view. Therefore, INJECTOR must prepare SDR with respect to this image. After calculating and storing SOF, EOF, CDIS, CMIS, and the other critical bytes, INJECTOR must divide image ② into block (BLK1, ... BLKN) and store the checksum numbers in the SDR table see fig. 3. Preparing the SDR header can be done easily if INJECTOR knows which algorithms and functions are used by LEP. Otherwise, INJECTOR must store the values of LOC1 and LOC2 in a message file so that LEP can prepare the SDR module header.

- Efficiency Analysis

The efficiency of SDS must be analyzed to see whether it can satisfy all of the requirements of IAVS or not. The efficiency of the SDS is determined according to the analysis criteria presented in [٤]. In what follows, SDS will be analyzed with respect to detection, prevention, eradication, and damage control.

1- Detection: SDR classified as a Target-Based detection.

a- VGS= Max.: Because SDR knows every thing about the protected executable, it can detect any change to this executable. Viruses cannot avoid detection by the SDR, because, they cannot infect the executable without changing its contents. Even stealth type viruses "see [7]" cannot deceive the SDR, because, they cannot get the opportunity to execute before the SDR. Therefore, SDR detection capability is independent from the virus generation date and VDC. Also, if LEP cannot find any one of SDR modules, it concludes that the protected executable was changed, and it might be infected by a virus.

b- TGS=1: Because SDR designed to detect viruses attached to the protected executable only.

c- Capability: The following is concluded from a & b: **"SDS can satisfy the requirement of ideal Target-Based detection".**

- Prevention: Both VEP-Based and VIP-Based prevention can be used in SDS.

- **VEP-Based Prevention:** SDR classified as a VEP-Based prevention anti-virus, because, it can detect the infection without giving viruses the opportunity to execute.

  a- VGS= Max.: SDR prevention is independent from the virus generation date and VDC.

  b- TGS= Max.: SDR will prevent the detected virus from infecting any target site within its environment.

c- Capability: The following is concluded from a, b, and the fact that the SDR is ideal with respect to detection: **"SDS can satisfy the requirement of ideal VEP-Based prevention".**

- **VIP-Based Prevention:** LEP classified as a VIP-Based prevention anti-virus, because, it can prevent the detected virus from infecting a new target site when executed in PHASE-TWO by using the virtual computer system.
  - **a- VGS:** It was mentioned in section 4.1 that the direct hardware access is the most vulnerable spot that the virus can use to penetrate the virtual computer system. Maximizing VGS thoroughly depends on the system that will implement the SDS. VGS can be maximized, if SDS is implemented on a computer system that provides hardware protection level (i.e. protected mode PC) with a permission or privilege level that will prevent an unauthorized program to direct access the hardware resources. Otherwise, VGS cannot be maximized.
  - **B- TGS= Max.:** LEP will try to prevent the detected virus from infecting any target site within its environment.
  - **c- Capability:** The following is concluded from a & b:
    - **"SDS can satisfy the requirement of ideal VIP-Based prevention if used on a system that provides hardware protection level (i.e. Protected Mode PC)".**
    - "**SDS cannot satisfy the requirement of ideal VIP-Based prevention if used on a system that lacks the hardware protection level (i.e. Real Mode PC)".**

- Eradication: As shown above, two eradication algorithms (FBEA and VFEA) are used by SDS:

- **FBEA:** FBEA uses Analysis-Based eradication technique. It analyzes the protected executable and eradicates any foreign block.
  - **a- VGS:** FBEA eradication capability depend on the virus type, shell or intrusive "see [7]", and NAB-distribution method (SBD or CBD). Therefore, the FBEA described above can eradicate only shell type viruses that use SBD method. Even though, FBEA can be upgraded to eradicate CBD viruses, the virus designers can design new strains of CBD-viruses that cannot be eradicated by FBEA. Therefore, the FBEA eradication capability depends on the virus generate date. This means that: **"FBEA cannot maximize VGS".**

NOTE: FBEA is a good enhancement in the design of eradication algorithms. It is very efficient eradication algorithm at the present time, because, it can eradicate, almost, all of the currently available viruses.

  - **b- TGS=1:** Because FBEA designed to eradicate viruses from the protected executable only.
  - **c- Capability:** The following is concluded from a & b: **"SDS cannot satisfy the requirement of ideal Analysis-Based eradication".**
- **VFEA:** VFEA uses Backup-Based eradication technique. In this case, the infected disk image represents the vulnerable copy, and the repaired (i.e. by the virus) memory image represents the backup copy.
  - **a- VGS:** VFEA can eradicate shell type viruses only and its eradication capability depend on the virus generation date, this means that: **"VFEA cannot maximize VGS".**
  - **b- TGS=1:** VFEA designed to eradicate viruses from the protected executable only.
  - **c- Capability:** The following is concluded from a & b: **"SDS cannot satisfy the requirement of ideal Backup-Based eradication".**

NOTE: VFEA is a good enhancement in the design of eradication algorithms. It is very efficient eradication algorithm at the present time, because, it can eradicate, almost, all of the currently available shell-type viruses.

- **Damage Control:** SDS system can use Virus-Based, Cell-Based, and Backup-Based damage control techniques.
  - **Virus-Based Damage Control:** In PHASE-ONE, SDR classified as a Virus-Based damage control anti-virus. Because it detects the virus before executing it. Clearly, the virus cannot cause any damage if not executed. Virus-Based damage control is similar to VEP-Based prevention, therefore, VGS= Max, TGS= Max, and: "**SDS can satisfy the requirement of ideal Virus-Based damage control".**
  - **Cell-Based Damage Control:** In PHASE-TWO, LEP classified as a Cell-Based damage control anti-virus, because, it tries to prevent viruses from destroying any target cell within its environment by using the virtual computer system. Clearly, direct hardware access is still a vulnerable spot. Cell-Based damage control is similar to VIP-Based prevention, therefore:
    - **"SDS can satisfy the requirement of ideal Cell-Based damage control, if used on a system with hardware protection level".**
    - **"SDS cannot satisfy the requirement of ideal Cell-Based damage control, if used on a system with no hardware protection level.**
  - **Backup-Based Damage Control:** In PHASE-TWO, LEP can store a backup copy of the virus target cells such as FAT, RD, BPS, and CMOS RAM before executing the virus. After the completion of the virus execution, LEP can repair any target cell that was destroyed by the virus, by using its backup. This technique is similar to saving the IVT before executing the virus and restoring it after the completion of the virus execution. In this case, LEP classified as a Backup-Based damage control anti-virus.

    a- VGS= Max.: LEP protection is independent from the virus generation date and VDC.

    b- TGS: TGS depend on the number of target cells selected by LEP.

    c- Capability: The following is concluded from a & b: LEP can protect any target cell in the selected T-group against destruction by any virus, therefore: **"SDS can satisfy the requirement of ideal Backup-Based damage control".**

- SDR Size Optimization

SDR will increase the protected executable size and the time needed to load and execute it. Therefore, minimizing the SDR size must be one of the design goals. One way to minimize the code of SDR is to implement the FBEA as a LEP (or OS) service. As mentioned earlier, the SDR calls the algorithm 'CHKS-ALG' with two arguments "Block number 'BLKi' and Checksum number 'CSN'". Clearly, it is not necessary to implement CHKS-ALG inside each SDR. Instead CHKS-ALG can be implemented as a global OS service that can be used by all SDRs. Note that the fact that the block size is different for different SDRs ensures that the general trend to avoid standard protection is not violated. Assuming that CHKS-ALG uses a standard CRC method, then, it can be implemented as a global service as follows:

CHKS-ALG

INPUT:

- BLKi= Start address of the block
- BLKS= Block Size in bytes
- CSN= The CRC checksum of the block.

OUTPUTS:

- C=0 ;Valid checksum
- C=1 ;Invalid checksum

In the same way, FBEA can be implemented as a global service. In this case, CHKS-ALG can be implemented inside the global FBEA. The standard Input/Output of the global FBEA can be defined as follows:

FBEA:

INPUTS:

- SMI, MIS, TOA, SOF, EOF
- N= Number of blocks in the memory image
- BLKS= Block Size
- Checksum Number Table: The table format is shown in Tab.9

OUTPUTS:

- C=0 ;Repair Succeed
- C=1 ;Repair Failed

- Conclusions and Discussion

A new approach in designing anti-virus system is presented in this paper. The proposed system is called Self-Defence System (SDS). The purpose of this research is to design an anti-virus system that complies with the requirements of ideal anti-virus system (IAVS) "see [4]. In [4], it was mentioned that Dependent-Defence System (DDS) couldn't satisfy the requirement of the IAVS because of four problems. Now let us consider these problems from the SDS point of view:

1- Lack of Knowledge: In principle, SDR must be implemented as integral part of the protected executable during the development process and it must know every thing about the protected executable. Therefore, SDS can satisfy the requirement of ideal detection. However, if the SDR injected inside an existing executable using vaccination, then, it is important to ensure that the executable is clean before injecting the SDR. If a virus exists in the executable prior to the injection, then, SDR cannot detect the presence of this virus.

2- Standard Protection: In principle, each SDR is designed and implemented by a different manufacturer. Therefore, the code and data of SDR will be different from one SDR to the other. For example, even though more than one SDR may use the FBEA, the block size (BLKS) selected by each SDR and the location of the checksum number table might be different. Clearly, viruses cannot deceive SDR without knowing this information.

3- Intended Vulnerability: In SDS, the SDR capability in coping with viruses is considered one of the executable quality factors. Good programs are those that can protect themselves against viruses efficiently. Therefore, any failure in the SDR operation and any undetectable infection have bad affects on the trustiness between the program developer and his customers. They may, simply, do not purchase his programs if they find that the SDRs associated with these programs are vulnerable.

4- Users Responsibility: In SDS, user responsibility is no longer a requirement or a factor of efficiency, because, the SDR designer cannot claim that his program cannot protect itself against the computer virus because the user do not use the protection program properly. SDR is embedded inside the protected program and its operation is, or must be, transparent for the computer user.

As a summary: **"SDS system solves the problems (Lack of knowledge, Standard Protection, Intended Vulnerability, and User Responsibility) found in DDS".**

Also, SDS provides a good support to design and use new sophisticated and efficient eradication algorithms. Even though FBEA is not ideal, it is considered a good enhancement in the design of Analysis-Based eradication programs, relative to the eradication method used in DDSs. In DDS, the eradication program can eradicate only 'Known' viruses. Any detected virus must be 'Identified' by security professionals, the identification information stored in a database, and a special eradication program designed to eradicate the virus. In SDS, FBEA eradicate any foreign block that exists between the protected executable blocks. FBEA doesn't care whether the foreign block belong to a known virus, unknown virus, logic bomb, time bomb, or a Trojan Horse.

Even though VFEA is not ideal, it is considered a good enhancement in the design of Backup-Based eradication programs, relative to the methods used by DDS. In DDS, for each target site (vulnerable copy), a backup copy must be stored in the disk. In SDS, VFEA depend on the virus to generate the backup-copy and then use it to repair the vulnerable copy, therefore, it has the following advantages relative to the method used by DDS:

- No disk space needed to store backups.
- There is no way that the virus will reside in the backup copy.
- There is no need to update the backup copy

References

-Garber, L., **"Antivirus Technology Offers New Cures"**, IEEE Computer Magazine, February, 1998.

-Pfleeger, C., P., **"SECURITY IN COMPUTING"**, Prentice-Hall Inc., 1989.

-Kaspersky, E., **"AVP Virus Encyclopedia"**, Version 1.3 (1992-1997).

-Hamid M. A. Abdul-Hussain, **"Computer Virology: Toward Designing An Ideal Anti-Virus System".** Engineering Journal, College of Engineering, University of Baghdad. Vol.8, No.3, 2002.

Tischer,M., **"PCINTERN SYSTEM PROGRAMING"**, Abacus, 1992

- عامر نزار فايز د. ،**"فيروسات الكمبيوتر"**، جامعة الإسراء-عمان-الأردن, 1995

- Hamid M. A. Abdul-Hussain & Hamed M. Shabib, **"Computer Virology: Formal Analysis of Computer Viruses".** Engineering Journal, College of Engineering, University of Baghdad. Vol.8, No.1, 2002.

List of Abbreviations

AT	Advanced Technology
BPS	Boot (or) Partition Sector
CBD	Complex Block Distribution
CDIS	Correct Disk Image Size
CMIS	Correct Memory Image Size
CRC	Cyclical Redundancy Check
CSN	CheckSum Number
DDS	Dependent Defense System
DIS	Disk Image Size
DOS	Disk Operating System
EBV	Executable Based Variable
EIF	Executable Infection Flag
EOF	End Offset
EPN	Executable Private Number
ERF	Executable Repair Flag
FAT	File Allocation Table
FBEA	Foreign Block Eradication Algorithm
IAVS	Ideal Anti-Virus System

IBM	International Business Machine
IIB	Image Information Block
IVT	Interrupt Vector Table
LBC	Loader Based Constant
LEP	Load and Execute Program
LOC	Trapdoor mark LOCATION
Max.	Maximum
Min.	Minimum
MIS	Memory Image Size
NAB	Number of Added Bytes
NCB	Number of Changed Bytes
OS	Operating System
PC	Personal Computer
RD	Route Directory
SBD	Simple Block Distribution
SDS	Self-Defense System
SDR	Self-Defense Routine
SIB	Secret Identification Block
SMI	Start of Memory Image
SOF	Start Offset
TGS	Target site/cell Group Size
TMV	Trapdoor Mark Value
TOA	Trapdoor Offset Address
VDC	Virus Deception Capability
VEP	Virus Execution Probability
VFEA	Virus Follower Eradication Algorithm
VGS	Virus Group Size
VIP	Virus Infection Probability

Tab.1

	LEP1		LEP2	
File Name	g	TMV=g+6	g	TMV= g⊕305
HAMED.COM	72+65+77=214	220	65-69= -4	-307
MUNTHER.EXE	77+85+78=240	246	85-84= 1	304
KARIM.SYS	75+65+82=222	228	65-73= -8	-311

Tab.2

Filename	LOC	[LOC]=TMV
HAMED.COM	500	503
MUNTHER.EXE	1024	1027
KARIM.SYS	856	859

Tab.3 Primary SDR Module Standard Format

LOC1+	Size(Byte)	Content	Function/Algorithm Name
00H	4	TMV11	f11(LOC1, LBC11)=g11(LOC1,LBC12)
04H	4	TMV12	f12(LOC1,LBC13)=g12(LOC1,LBC14)
08H	4	CLEN	f13(LEN)
0CH	4	CSN	CHKS-ALG1
10H	LEN	SDR	SDR code/data

Where:

LEN= SDR Module Length In bytes
CLEN= Coded Length
CSN= CheckSum Number
CHKS-ALG= CheckSum Algorithm

Tab.4 Redundant SDR Module Standard Format

LOC2+	Size(byte)	Content	Function/Algorithm Name
00H	4	TMV21	f21(LOC2,LBC21)=g21(LOC2,LBC22)
04H	4	TMV22	f22(LOC2,LBC23)=g22(LOC2,LBC24)
08H	4	CLEN	f23(LEN)
0CH	4	CSN	CHKS-ALG2
10H	LEN	SDR	SDR code/data

Tab.5 Message File Entry Format		
Offset	Size (Byte)	Content
00H	11	Executable Name (E1, E2, ...)
0BH	4	Value of LOC1
0FH	4	Value of LOC2

Tab.6 IIB Standard Format		
Offset	Contents	Description
00H	SMI	Start Address of the executable Memory Image
04H	MIS	Memory Image Size
08H	DIS	Disk Image Size
0CH	EPN	Executable Private Number
0EH	TOA	The SDR Trapdoor Offset Address relative to SMI
12H	LEP-TRAP	Far pointer to the trapdoor of LEP.

Tab.7	
Conditions	Description
(TOA=SOF) AND (MIS-TOA=EOF)	Clean executable
(TOA>SOF) AND (MIS-TOA=EOF)	Virus block before the SDR
(TOA=SOF) AND (MIS-TOA>EOF)	Virus block after the SDR
(TOA>SOF)AND (MIS-TOA>EOF)	CBD Virus

Tab.8 SIB Standard Format		
Address	Size (Byte)	Content
00	2	Length of SIB in bytes
02	?	SIB contents

Tab.9 “SDR Checksum Numbers Table”	
Block Number	CRC-Checksum (CSN)
1	CSN1
2	CSN2
.	.
.	.
N	CSNn

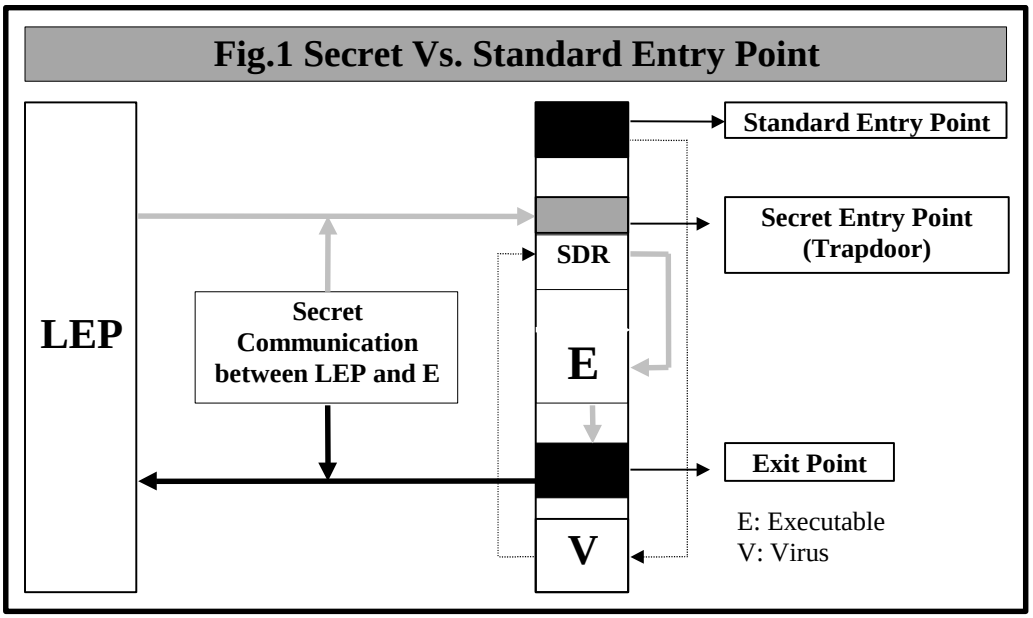


Fig.1 Secret Vs. Standard Entry Point
LEP
Secret Communication between LEP and E
SDR
E
V
Standard Entry Point
Secret Entry Point (Trapdoor)
Exit Point
E: Executable
V: Virus

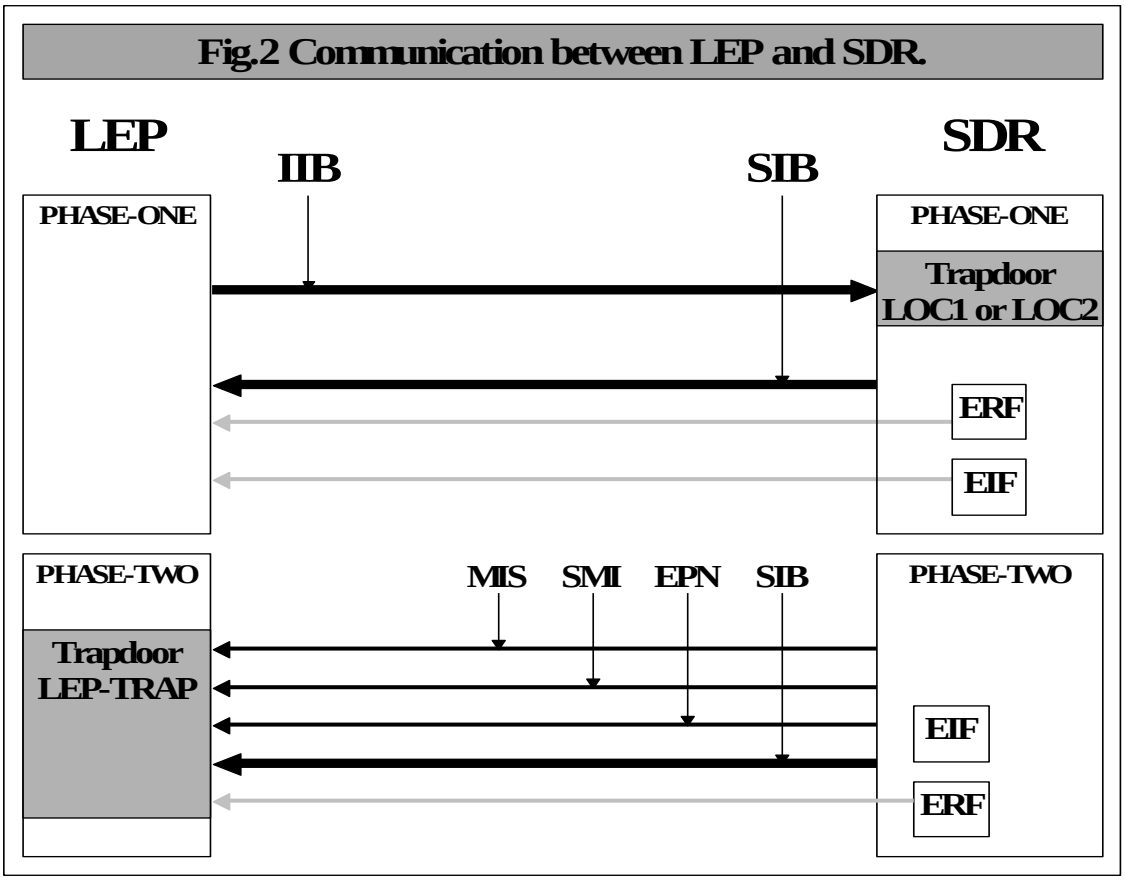


Fig.2 Communication between LEP and SDR.
LEP
SDR
IIB
SIB
PHASE-ONE
PHASE-ONE
Trapdoor LOC1 or LOC2
ERF
EIF
PHASE-TWO
PHASE-TWO
MIS
SMI
EPN
SIB
Trapdoor LEP-TRAP
EIF
ERF

Fig.3 The Protected Executable In Memory

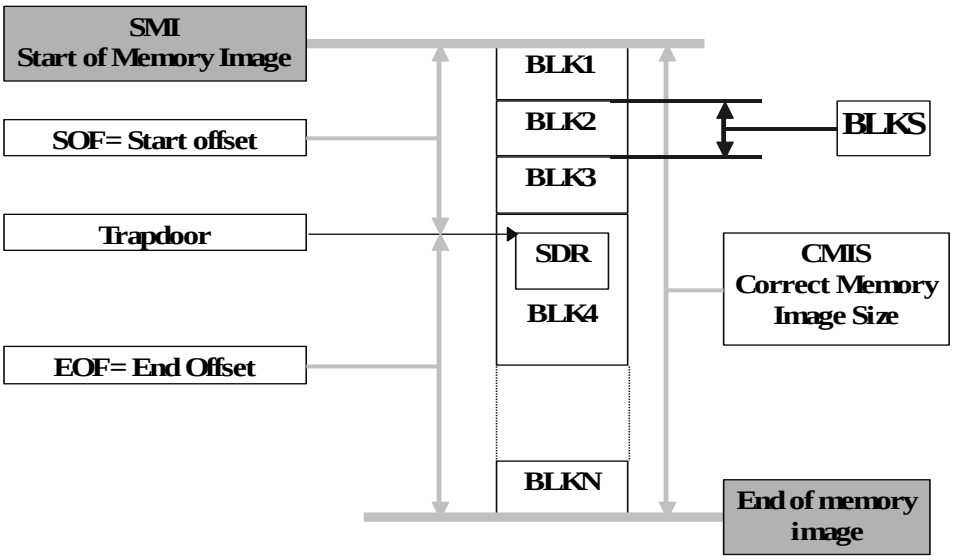


Fig.4 Simple Virus Distribution

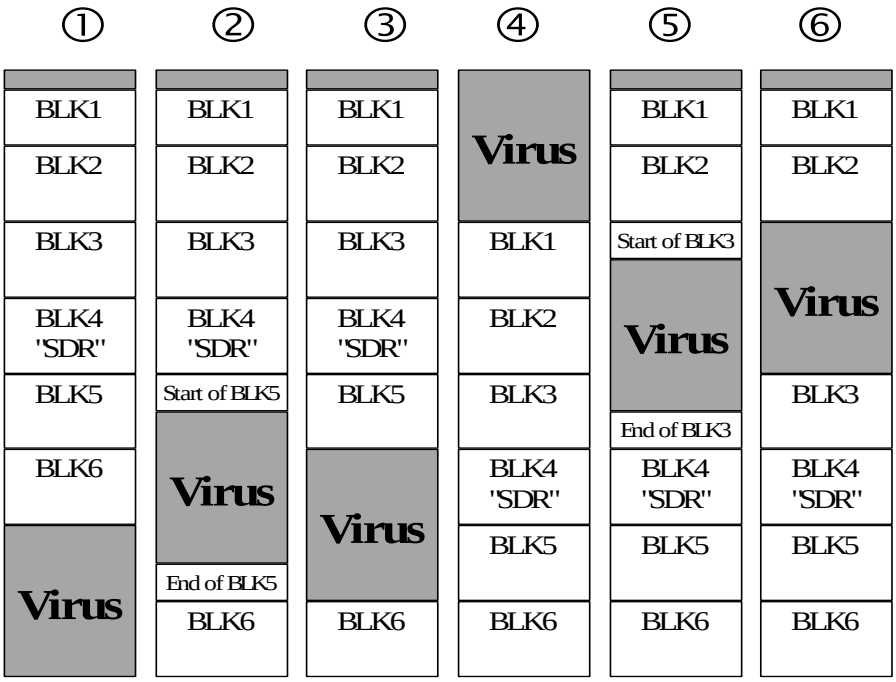


Fig.5 Determining S and E

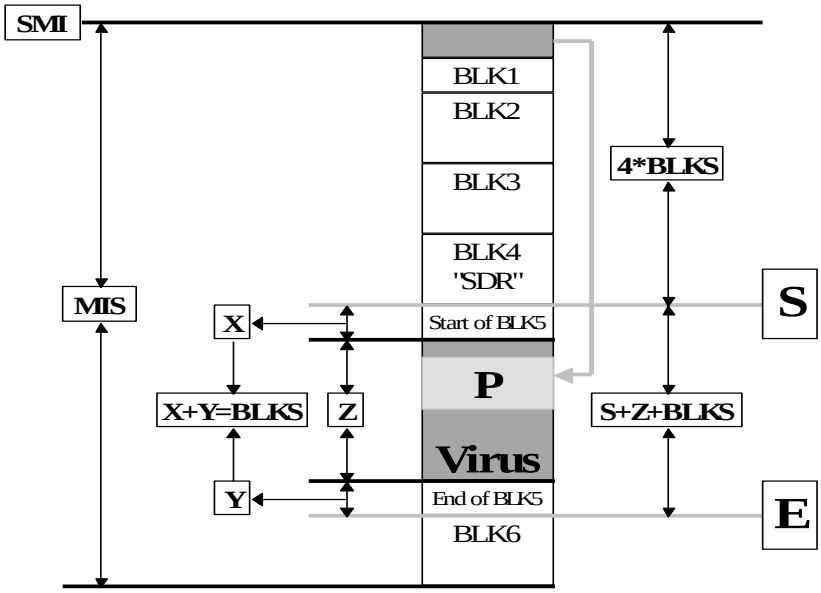
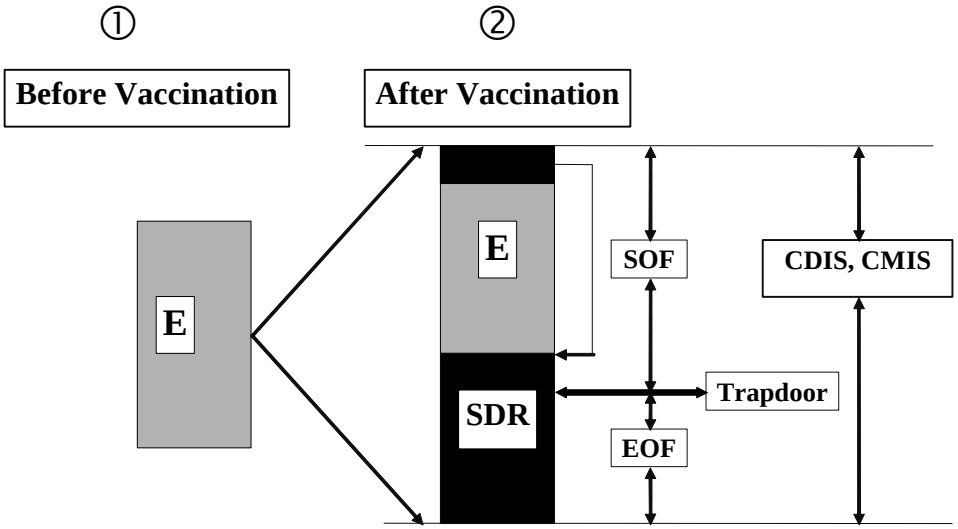


Fig.6 The Executable E Before and After Vaccination.



EQUILIBRIUM ADSORPTION OF METHANE AND CARBON DIOXIDE ON 5A MOLECULAR SIEVE

Dr. Abbas Khalaf Mohammed Shua'ab
Chemical Engineering Department, Baghdad University, P. O. Box 47024, Baghdad, Iraq

ABSTRACT

Experimental and theoretical studies have been carried out to study the adsorption of methane and carbon dioxide on 5A molecular sieve. Adsorption equilibrium isotherms of methane and carbon dioxide are reported for the temperature range 303–333 K and pressure up to 2.5MPa. Experimental data were obtained using a static system for gas–solid adsorption. The Langmuir adsorption equilibrium equation gave good predictions. Adsorption of methane and carbon dioxide on 5A molecular sieve is purely physical since the isosteric heat of adsorption was found to be equal to 14.804 and 37.218 kJ mole^{-1} for methane and carbon dioxide, respectively.

الخلاصة

تم الحصول على توازن الامتزاز الايزوثرمي لغاز الميثان و غاز ثاني اوكسيد الكربون باستخدام المناخل ألجزيئيه نوع 5A و عند درجات حرارة تتراوح من ٣٠٣كلفن إلى ٣٣٣كلفن و ضغط يصل إلى ٢,٥ميكاباسكال. النتائج ألمختبريه تتطابق بشكل جيد مع النموذج التجريبي للانكماير. امتزاز غاز الميثان و غاز ثاني اوكسيد الكربون على المناخل ألجزيئيه نوع 5A هو امتزاز فيزياوي وذلك بسبب إن حرارة الامتزاز لغاز الميثان تساوي ١٤,٨٠٤كيلوجول/مول و لغاز ثاني اوكسيد الكربون تساوي ٣٧,٢١٨كيلوجول/مول.

KEYWORDS
molecular sieve; adsorption; methane; carbon dioxide)

INTRODUCTION

The increasing importance of adsorption processes for the separation of gaseous mixtures is due to the high selectivity and adsorption capacity of solid adsorbents, the less extreme operation conditions needed, and the small energy consumption required. Adsorption is becoming a competitive operation that can advantageously substitute for other separation operations such as distillation or absorption. The advantages are especially attractive when the problem is the separation of light gases, since their separation by distillation or absorption requires expensive high pressure units.

The adsorptive separation is achieved by one of the three mechanisms: steric, kinetic or equilibrium effect. The steric effect derives from the molecular sieving property of zeolites. In this case only small and properly shaped molecules can diffuse into the adsorbent, whereas other molecules are totally excluded. Kinetic separation is achieved by virtue of the difference in diffusion rates of different molecules into the adsorbent. By far, however, most processes operate through the equilibrium (or competitive) adsorption of the mixture, and hence are called equilibrium separation processes. Commercial separation processes using molecular sieve zeolites are mainly equilibrium processes (Kapoor & Yang, 1989).

Separation of bulk gas mixtures containing methane and carbon dioxide is an important problem in the chemical industry. Many natural gas reservoirs contain up to 50% CO_2 as impurity. The effluent gas from an oil well undergoing CO_2 flooding for enhanced oil recovery may contain 20–80% CO_2 and CH_4. It may be necessary to separate the CH_4 and the CO_2 from these gases in order to improve their fuel value and to recover and reuse the CO_2.

Data on the adsorption of methane and carbon dioxide on molecular sieves are scarce in the literature. Some low–pressure data and correlations have been reported for Linde 5A and 13X molecular sieve, where carbon dioxide was found to be preferentially adsorbed on the former molecular sieve (Tańczyk & Warmuziński, 1998; Kumar, 1989; Hyun & Danner, 1982). Limitation are placed on the usefulness of these data for the design of commercial molecular sieve–based separation systems, since a large part of these data was obtained at temperatures and pressures which are not ordinarily considered practical from a commercial stand point.

This paper reports a fundamental equilibrium adsorption data obtained for methane and carbon dioxide on a single sample of 5A molecular sieve (supplied by Rhone Poulenc industries). To simulate industrial conditions, the temperature ranged from ambient to 60 ˚C and the pressure to 2.5 MPa. These data are of value in process and design studies of 5A molecular sieve–based separation systems involving the separation of carbon dioxide from natural gas and oil well effluent.

EXPERIMENTAL WORK

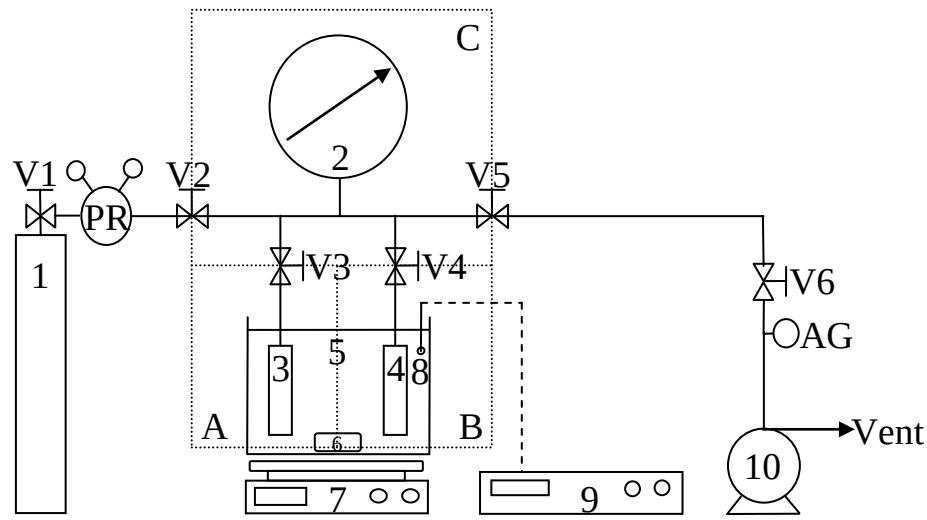
A schematic diagram of the apparatus is shown in figure 1. All tubing was thick wall stainless steel, to minimize the dead volume. The tubing was connected by Crawford Swagelock compression fittings. Parker CPI series severe service union bonnet valves, with a hardened ball stem tip, were chosen because of their durability under repetitive use. A metal stem tip was selected instead of a soft stem tip to avert destruction of the valve seat in the event that particulate matter should escape from the adsorption chamber, enter the valve and score the valve seat.

The adsorption chamber and reservoir were 20 cm^3 iron sample cylinders. Entrainment of adsorbent particles in the tubing and valving system was prevented by employing a dual screen method. This was accomplished by inserting Pyrex glass wool and a conical brass wire mesh screen, capable of containing particles.

The gas pressure was measured via a Heise bourdon gauge with a 430 mm diameter dial and a pressure range from 0 to 6891 kPa. The accuracy of the gauge was rated at 0.1% of the full scale and the sensitivity at 0.02% of the full scale at all points. The pressure gauge enclosed a dead volume of approximately 30 cm^3.

A water bath was employed to provide a constant temperature environment up to 353 K for both the adsorber and the reservoir. The bath was thermo–stated and vigorously mixed by using a magnetic stirrer thermo–stat hot plate. The temperature of the bath was continuously measured and recorded via a digital recorder and a thermocouple wire, calibrated with mercury thermometer.

The whole equipment was evacuated by vacuum pump and absolute gauge pressure prior to each experiment. The feed gas supplied from cylinder was regulated by pressure regulator.\



1	Gas cylinder	6	Stir bar	V	Manual valve
2	Heise pressure gauge	7	Magnetic stirrer thermostat hot plat	PR	Pressure regulator
3	Reservoir bomb	8	Chromel–alumel thermocouple	AG	Absolute gauge
4	Adsorber bomb	9	Digital recorder		
5	Water bath	10	Vacuum pump		

Fig (1) Schematic diagram of apparatus used for adsorption equilibrium measurement

The volume of the reservoir and the adsorber (together with the Swagelock fittings) and the other sections of the system were measured manometrically using nitrogen gas. Section C was charged with nitrogen gas to a known pressure P_1 = 2.5 MPa. Valve 5 was connected to a standard known volume previously evacuated cylinder vessel V_1 = 1 liter and opened until equilibrium was reached with a new pressure P_2, then the volume of section C was determined by the ideal gas equation of state:

$$V_C = \frac{P_2 V_1}{P_1 - P_2} \qquad (1)$$

The volume of section A and B were determined in the same manner.

The granular adsorbent (10 g) was packed into the adsorption vessel and degassed for overnight under vacuum at a pressure of less than 3 Pa to clean the adsorbent. This

pretreatment enabled us to make adsorption measurements without changing the adsorbent. After the adsorbent had been regenerated, the vacuum pump was switched off. The water bath with electrical stirrer thermo–stated hot plate was switched on until constant temperature was reached.

The procedure for measuring equilibrium adsorption was as follow:

1. Section A and C were pressurized with the adsorbate, corresponding to a pressure P_3 then the number of moles of the adsorbate was estimated according to the relation:

$$n_1 = \frac{P_3(V_C + V_A)}{RT} \tag{2}$$

Passage of gas to section B was allowed until equilibrium ($\Delta P \leq 3$ kPa/h). The equilibrium pressure was recorded as P_4.

2. Adsorber was isolated from the system. The number of moles of the adsorbate was estimated according to the relation:

$$n_2 = \frac{P_4(V_C + V_A + V_B)}{RT} \tag{3}$$

The amount adsorbed of the adsorbate at partial pressure equal to P_4 is determined according to a material balance equation:

$$q = \frac{n_1 - n_2}{W} \tag{4}$$

RESULTS & DISCUSSIONS

The partial pressure and temperature ranges were chosen for each component in viewing to the expected values in the natural gas and oil well effluent properties and component vapor pressure, which almost must be greater than the component partial pressure under specified temperature, to avoid condensation of gases on the adsorbent.

The experimental results for single gas equilibrium were fitted with Langmuir and Freundlich equations. BET and other equations that represent multi–layer adsorption were excluded because it is unlikely that multi–layer adsorption could have occurred in this study were all temperatures above the critical (Yang & Saunders, 1985).

The Langmuir equation takes the form:

$$\frac{q}{q_m} = \frac{BP}{1+BP} \tag{5}$$

The coefficients B and q_m depend on temperature in accordance to the following equations:

$$q_m = a_1 T^{a_2} \tag{6}$$

$$B = a_3 e^{a_4/T} \tag{7}$$

The Freundlich equation takes the form:

$$q = k_F P^{n_F} \tag{8}$$

The coefficients k_F and n_F depend on temperature in accordance to the following equations:

$$k_F = b_1 T^{b_2} \tag{9}$$

$$n_F = b_3 + b_4/T \tag{10}$$

The objective function (average relative error) used in the minimization routine was defined as:

$$E = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\left(\frac{q_{i,exp} - q_{i,cal}}{q_{i,exp}}\right)^2} \qquad (11)$$

where N is the number of experimental points and $q_{i,exp}$ and $q_{i,cal}$ are the experimental and calculated values of q_i, respectively.

The calculated constants for the two model equations along with the average relative error values were presented in table 1. The table shows that Langmuir equation correlates the experimental data with mean average relative error of 1.59% while 2.29% for Freundlich equation. Therefore, the best fit was achieved with Langmuir equation. The experimental equilibrium data was correlated by Langmuir equation and presented in figures 2 and 3.

Table (1) Summary of single gas adsorption isotherms results

Adsorbate	a_1	a_2	a_3	a_4	E %
CO_2	4.71×10^{-1}	-0.86	2.27×10^{-4}	999.67	2.67
CH_4	8.25×10^{-1}	-0.99	3.24×10^{-7}	705.96	0.51

Adsorbate	b_1	b_2	b_3	b_4	E %
CO_2	2.99×10^{-3}	-0.19	3.12×10^{-2}	180.99	2.80
CH_4	4.21×10^{-44}	15.42	-8.7×10^{-1}	385.45	1.77

All curves for any specified gas are similar in shape and are of the classic adsorption isotherm form. However, it was concluded that the strength of adsorption for CO_2 is greater than for CH_4, since it was recognized that CO_2 has isotherms with greater slopes at $P \rightarrow 0$, which indicates that the affinity for adsorption of CO_2 is greater than for CH_4.

The results also indicate that there is a fair correlation for the amount of adsorption on an adsorbent correspond to increased adsorption with higher gas molecular weight or critical temperature. This result is in agreement with published works (Lewis et al., 1950). Also, vapor pressure can be expected to be the predominant factor favoring adsorption. The much less volatile gas is more likely to condense on the surface since it has a much larger relative saturation value (P / P^*) than the more volatile one. This result was verified by several authors (Hyun & Danner, 1982).

Adsorption isotherms for the CH_4 and CO_2 gases on 5A molecular sieve at moderate temperature ranges are presented in figures 2 and 3. The marked decrease in adsorption with increasing temperature is obvious from these diagrams.

The heat of adsorption (ΔH) was calculated by application of the Clausius–Clapeyron equation (Smith, 1988) at neighboring temperatures. When this equation is applied to the two phase system of gas and adsorbed component on the surface, we get:

$$\left(\frac{dP}{dT}\right)_{\theta} = \frac{\Delta H}{T(V - V_a)} \tag{12}$$

where V and V_a are the volumes per mole of adsorbed component in the gas and on the surface, respectively. V_a is very small and could be neglected. By assuming the ideal gas law for V we get:

$$\Delta H = -R\left(\frac{d\ln P}{d\,1/T}\right)_{\theta} \tag{13}$$

Table 2 shows the results of this equation.

Table (2) Heat of adsorption and condensation for gases

Component	ΔH, (kJ mole^{-1})	ΔH_{cond} at BP (kJ mole^{-1})
CH_4	14.804	8.180
CO_2	37.218	Sublimes at 195 K

In both cases the heats of adsorption are higher or expected to be higher than the latent heats of condensation but are low enough to be characterized as physical adsorption.

CONCLUSIONS

1. Equilibrium isotherms for single component adsorption of methane and carbon dioxide on 5A molecular sieve can be correlated by Langmuir equation.
2. The strength of adsorption for carbon dioxide is greater than for methane on 5A molecular sieve.

3. The heat of adsorption for both methane and carbon dioxide on 5A molecular sieve always greater than the latent heat of condensation.

NOMENCLATURE

$a_1,.., a_4$	Langmuir isotherm coefficients	-
$b_1,.., b_4$	Freundlich isotherm coefficients	-
k_F	Constant in Freundlich equation	Pa^{1-nF}
n	Number of moles	mole
n_F	Parameter in Freundlich equation	-
q	Adsorbed phase concentration	mole g^{-1}
y	Mole fraction in gas phase	-
B	Adsorption coefficient in Langmuir equation	Pa^{-1}
E	Average relative error	-
N	Number of experimental points	-
P	Pressure	Pa
P^*	Vapor pressure	Pa
R	Ideal gas constant	J $mole^{-1}$ K^{-1}
T	Temperature	K
V	Volume	m^3
W	Mass of adsorbent	g
θ	Fraction of surface coverage	-
ΔH	Heat of adsorption	J $mole^{-1}$

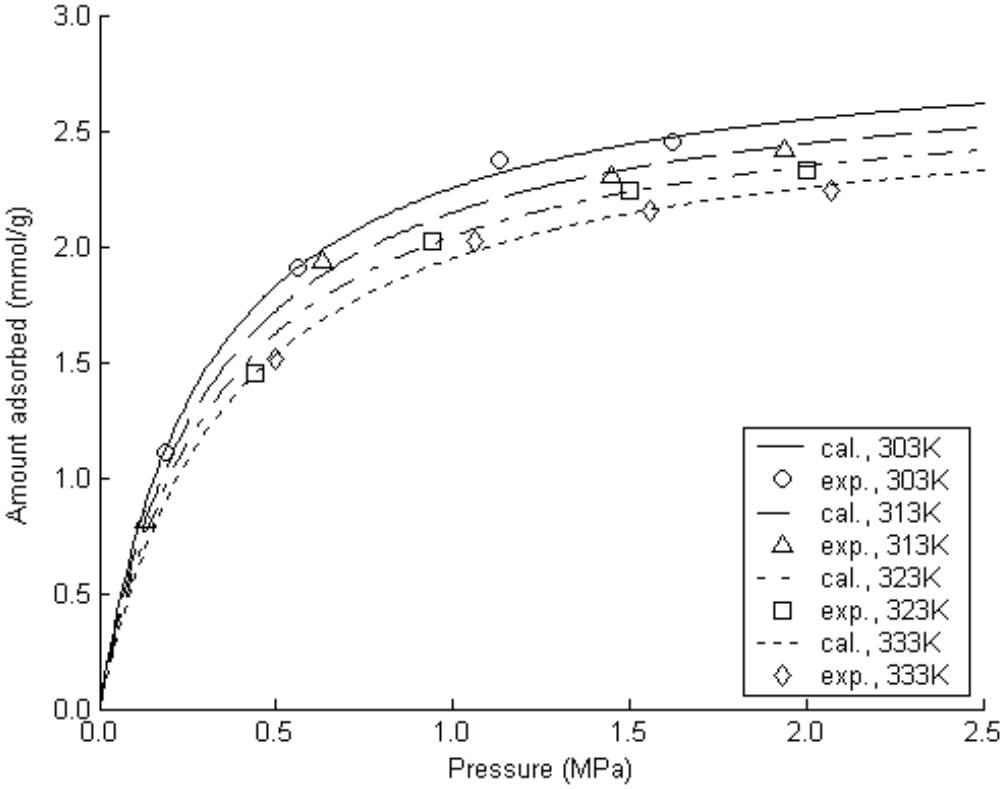


Figure (2) Adsorption equilibrium isotherms for the system Methane-5A Molecular sieve

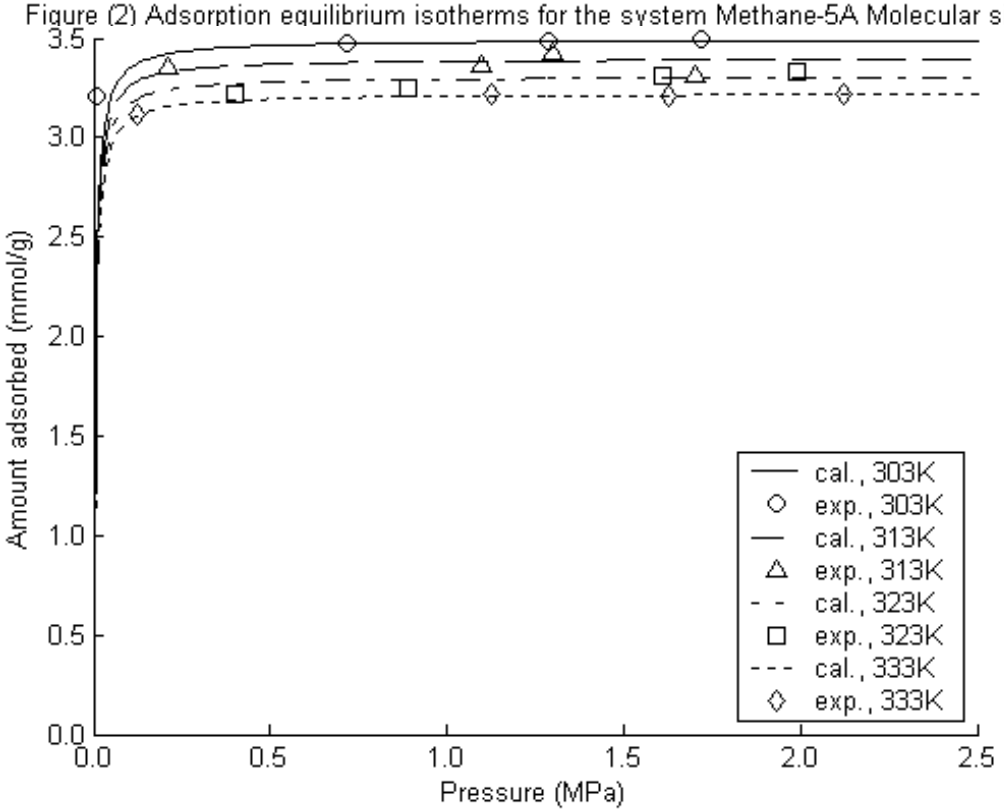


Figure (3) Adsorption equilibrium isotherms for the system Carbon dioxide-5A Molecular sieve correlated with Langmuir equation

REFERENCES

- Hyun S.H. & Danner R.P., Equilibrium adsorption of ethane, ethylene, isobutane, carbon dioxide and their binary mixtures on 13 X molecular sieve, *J. Chem. Eng. Data*, **27**, 196 (1982).
- Kapoor A. & Yang R.T., Kinetic separation of methane-carbon dioxide mixture by adsorption on molecular sieve carbon, *Chem. Eng. Sci.*, **44**, 1723 (1989).
- Kumar R., Adsorption column blow down, *Ind. Eng. Chem. Res.*, **28**, 1677 (1989).
- Lewis W.K., Gilliland E.R., Chertow B. & Cadogan W.P., Adsorption equilibria, *Ind. Eng. Chem.*, **42**, 1326 (1950).
- Smith J.M., *Chemical engineering kinetics*, 3rd ed., McGraw–Hill, Singapore, 1988.
- Tańczyk M. & Warmuziński K., Multi-component pressure swing adsorption. Part II. Experimental verification of the model, *Chem. Eng. Proc.*, **37**, 301 (1998).
- Yang R.T. & Saunders J.T., Adsorption of gases on coals and heat-treated coals at elevated temperature and pressure, *Fuel*, **64**, 616 (1985).

DESIGN CHARTS FOR MACHINE FOUNDATIONS

Mohammed Yousif Fattah
Assistant Professor, Dept. of Building
Construction Engineering, University of
Technology, Iraq.

Ahmed A. Al-Azal Al-Mufty
Assistant Professor, Dept. of Civil
Engineering, University of Baghdad, Iraq

Hula Taher Al- Badri
Formerly graduate student, Dept. of Civi
Engineering, University of Baghdad, Iraq.

ABSTRACT
The problems of design of machine foundations for the special case of vertical mode of vibration for block foundation are presented in this paper. The empirical design method is used to get the results using a computer program **MATHCAD** dealing with the parameters related to the machine. Design charts that are prepared to be a guide for the designer engineer are drawn. The design charts are based on the variables limitations including the properties of the soil, machine and foundation. The design charts are based on three displacements which are acceptable for design of the machine foundation.

الخلاصة
يتضمن البحث دراسة المشاكل التي تتعرض لها اسس المكائن بالاتجاه العمودي للاسس الكتلية . لقد استخدمت الطريقة الوضعية للحصول على النتائج بمساعدة برنامج حسابي (**MATHCAD**) والذي يتعامل مع حدود ترتبط مع الماكنه . لقد تم اعتماد جداول للتصميم كمرشد للمهندس المصمم في الموقع . جداول التصاميم اعتمدت محددات المتغيرات والتي تتضمن خواص كل من التربة، الماكنة والاساس. حيث اعتمدت على ثلاث ازاحات والتي تعتبر مقبولة من ناحية التصميم.

KEY WORDS: Machine foundation, Design charts, Empirical methods.

INTRODUCTION

The design of machine foundations is a trail-and-error procedure involving three interrelated steps (Gazetas and Roesset, 1979):
1) Establishment of desired foundation performance (design criteria),
2) Determination of magnitude and characteristics of the dynamic loading,
3) Estimation of anticipated translational and rotational motion of machine-foundation-soil system.

The design of a machine foundation is more complex than that of a foundation which supports only static loads. In machine foundations, the designer must consider, in addition to the static loads,

the dynamic forces caused by the working of the machine operation. These dynamic forces are, inturn, transmitted to the foundation supporting the machine (Srinivasulu and Vaidyanathan, 1976).

DESIGN LIMITS OF MACHINE FOUNDATION FOR EMPIRICAL METHODS

The design of block foundation for centrifugal or reciprocating machine starts with preliminary sizing of the block, which has been found to result in acceptable configuration as (Arya et al., 1979):

1. The bottom of block foundation should be above water table. It should not be resting on back filled soil nor on a special sensitive soil.
2. The mass of rigid foundation equals (2-3) times the mass of supported machine (for centrifugal), while the mass of rigid foundation equals (3-5) times the mass of supported machine (for reciprocating).
3. The top of block is usually kept (0.3 m) above finished floor or pavement elevation to prevent damage from surface water run off.
4. The vertical thickness of block should not be less than (0.61 m). The thickness seldom less than one-fifth the least dimension or one-tenth the largest dimension.
5. The foundation should be wide enough to increase damping in the rocking mode. The width should be at least (1-1.5) times the vertical distance from the base to machine centerline.
6. The combined center of gravity should coincide with the center of gravity of the foundation.
7. For large reciprocating machines, it may be desirable to increase the embedded depth in soil such that 50% to 80% of the depth, this will increase the lateral restrain and damping ratio for all modes of vibration.
8. Static bearing capacity q_{all}: proportion of footing area for 50% of allowable soil pressure, which means that the actual soil pressure should be less than 50% of static bearing capacity q_{all}. The actual soil pressure equals to the weight of machine and foundation divided by the base area of footing as shown:

   Actual soil pressure = $\dfrac{W_{mach.} + W_{fou.}}{L_f \cdot B_f}$
9. Static settlement must be uniform; center of gravity of footing and machine load should be within 5% of each linear dimension from the foundation center.
10. Bearing capacity: static plus dynamic loads. The sum of static and modified dynamic loads should not create bearing pressure greater than 75% of allowable soil pressure given for static load condition q_{all}.
11. The magnification factor (M) should preferably be less than (1.5). The magnification factor can be defined as the ratio of dynamic displacement to the static displacement as shown in **Table (1)**.
12. Vibration amplitude (Y), at operating frequency is shown in **Fig. (1)**. The maximum amplitude of motion for the foundation system should lie in zones A or B.
13. The velocity which equals ($2\pi f$ x displacement amplitude) compares with the limiting value in **Table (2)** and **Fig. (1)**.
14. The acceleration which equals ($4\pi^2 f^2$ x displacement amplitude) should be tested for zone B in **Fig. (1)**.

    where:

f = Operating speed of machine = $\dfrac{\omega}{2\pi}$

15. Resonance: the acting frequencies of machine should have at least a difference of ±20 % with the resonance frequency of **Table (1)**.

$$0.8\, fmr \geq f \geq 1.2\, fmr$$

16. The horizontal translation and the rocking mode needs not be coupled if:

$$\sqrt{f^2{}_{nx} + f_{n\psi}{}^2} \Big/ f_{nx} \quad \left(f_{n\psi} \leq 2/3\ f\right)$$

where:

$f^2{}_{nx}$ = natural frequency in the x- direction, rpm.

$f_{n\psi}$ = natural frequency in the rocking direction, rpm.

Table (1) – Summary of derived expressions for a single-degree-of-freedom system (Arya et al., 1979).

Expression	Constant Force Excitation F_0 Constant	Rotating Mass-type Excitation $F_0 = m_i\, e\, \omega^2$
Magnification factor	$M = \dfrac{1}{\sqrt{(1-r^2)+(2Dr^2)}}$	$M = \dfrac{r^2}{\sqrt{(1-r^2)+(2Dr^2)}}$
Amplitude frequency f	$Y = M(F_o/k)$	$Y = M_r(m_i e/m)$
Resonance frequency	$f_{mr} = f_n\sqrt{1-2D^2}$	$f_{mr} = \dfrac{f_n}{\sqrt{1-2D^2}}$
Amplitude at resonance frequency f_r	$Y_{max} = \dfrac{F_o/\kappa}{2D\sqrt{1-D^2}}$	$Y_{max} = \dfrac{(m_i e/m)}{2D\sqrt{1-D^2}}$
Transmissibility factor	$T_r = \dfrac{\sqrt{1+(2Dr^2)}}{\sqrt{(1-r^2)+(2Dr^2)}}$	$\overline{T}_r = \dfrac{r^2\sqrt{1+(2Dr^2)}}{\sqrt{(1-r^2)+(2Dr^2)}}$

where:

r = ω/ω_n

ω_n = Natural circular frequency rad / sec.

ω = Frequency of excitation force = $\sqrt{(k/m)}$, rad / sec.

k = Spring constant, kN /m

m = Mass of machine and foundation, kg

m_i = Rotating mass, kg

D = Damping ratio = C/C_c

C = Damping

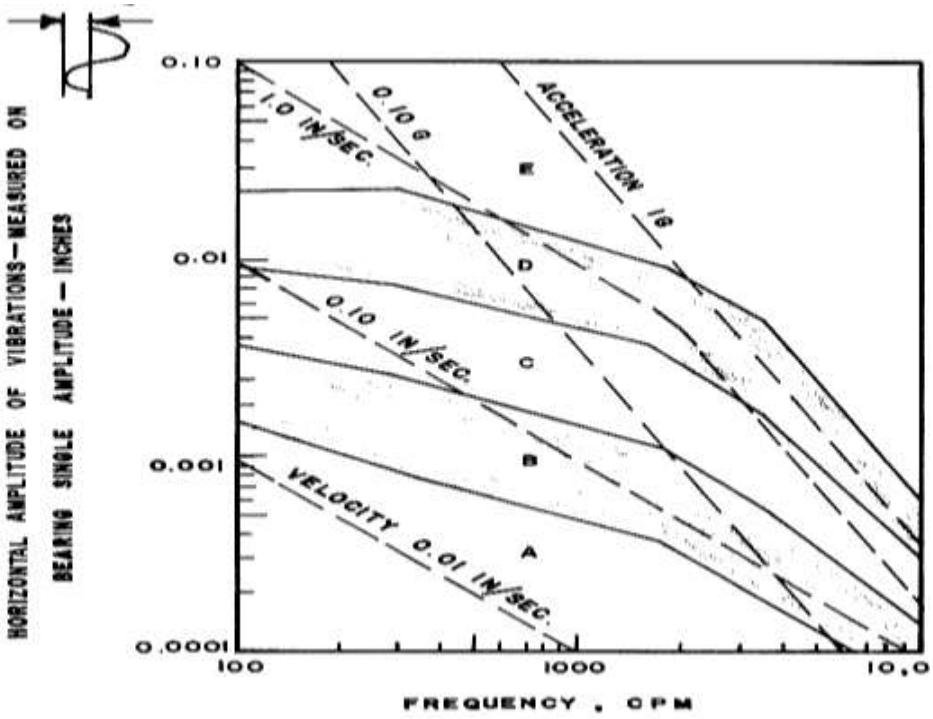
C_c = Critical damping = $2\sqrt{km}$

e = Eccentricity of unbalance mass to axis of rotation at operating speed, m

f_n = Natural frequency, rpm

f_{mr} = Resonant frequency for rotating mass-type excitation, rpm

M = Dynamic magnification factor
M_r =Magnification factor
F_0 = Amplitude of excitation force, kN
T_r = Force transmitted / F_0
$\overline{T_r}$ = Force transmitted / m_i e ω^2_n
Y = Amplitude at frequency *f*



Horizontal Peak Velocity (m/sec.)	Machine Operation
<0.00013	Extremely smooth
0.00013-0.00025	Very smooth
0.00025-0.00051	Smooth
0.00051-0.00101	Very good
0.00101-0.00203	Good

0.00203-0.00406	Fair
0.00406-0.008	Slightly rough
0.008-0.016	Rough
>0.016	Very rough

Fig. (1): Vibration performance of rotating machines (Harr, 1966).

A No faults. Typical new equipment.
B Minor faults. Correction wasted dollars.
C Faulty. Correction within 10 days to save maintenance dollars.
D Failure is near. Correct within two days to avoid breakdown.
E Dangerous. Shut it down now to avoid danger.

Table (2) –General machinery-vibration-severity data (Richart et al., 1970).

FORMULATION OF THE PROBLEM

The objective is to provide a clear image of design for machine foundation by using empirical methods. The empirical method, which is dependent on the theory of elastic half-space, the parameters of machine foundation and soil required for analysis are first obtained.

In this theory the footing is assumed to rest on the surface of the elastic half space and to have simple geometrical areas of contact, usually circular, but other shapes such as rectangular or long strip are possible (Arya et al., 1979). This theory includes the dissipation of energy throughout the half-space by "geometric damping" and allows calculation of finite amplitude of vibration at the "resonant frequency". The method is an analytical procedure, which provides a rational means of evaluating the spring and damping constants for incorporation into lumped-parameter, mass-spring-dashpot-vibrating systems.

The parameters of machine include the weight of machine depending on its type, which may be reciprocating compressor that is relatively heavy machine and generate vibrating forces of substantial magnitude at low operating frequency. It is also important to know the primary and secondary compressor speed in (rpm) and the primary and secondary forces and moments.

The parameters of soil on which the footing is assumed to rest on are obtained considering the surface of elastic half space and to have simple contact area. For the present case, the footing is rectangular with dimensions of $L_f \times B_f \times h$ (depending on limits or experience of the designer).

The type of soil is also considered, which is in this problem silty sand gravel (medium dense) including the density of soil (γ), shear modulus (G) and Poisson's ratio (ν). The allowable bearing capacity q_{all} and the permanent settlement of the soil (S_{tt}) are also considered.

EQUATIONS OF THE MACHINE FOUNDATION

The foundation of machine when designed requires knowledge of the dimensions for design; these dimensions are supplied by the manufacturer of the machine or depending on the experience of the designer. The dimensions of the foundation are considered as (L_f, B_f, h) in which the weight of the foundation equals to:

$$W_{foun} = L_f \times B_f \times h \times \gamma_c$$

where:

γ_c = the unit weight of concrete = 23.5 kN/m^3

The effect of the shape of foundation is approximately considered by equivalent radius (r_o). So for rectangular foundation, the equivalent radius is:

$$r_{0z} = \sqrt{\frac{B_f \times L_f}{\pi}} \qquad (1)$$

To calculate the equivalent spring constant for the vertical direction, the spring constant embedment factor in vertical direction and the spring coefficient have to be specified as follows:

$$\eta_z = 1 + 0.6(1-\nu) \cdot \frac{h_0}{r_{oz}} \qquad (2)$$

where: h_0 is the effective depth of embedment of the foundation.

The spring coefficient for vertical direction (β_z) is obtained from **Fig. (2)** (Srinivasulu and Vaidyanathan, 1976) as below:

$$\kappa_z = \frac{G}{1-\upsilon} \beta_z \sqrt{L_f \times B_f} \ \eta_z \qquad (3)$$

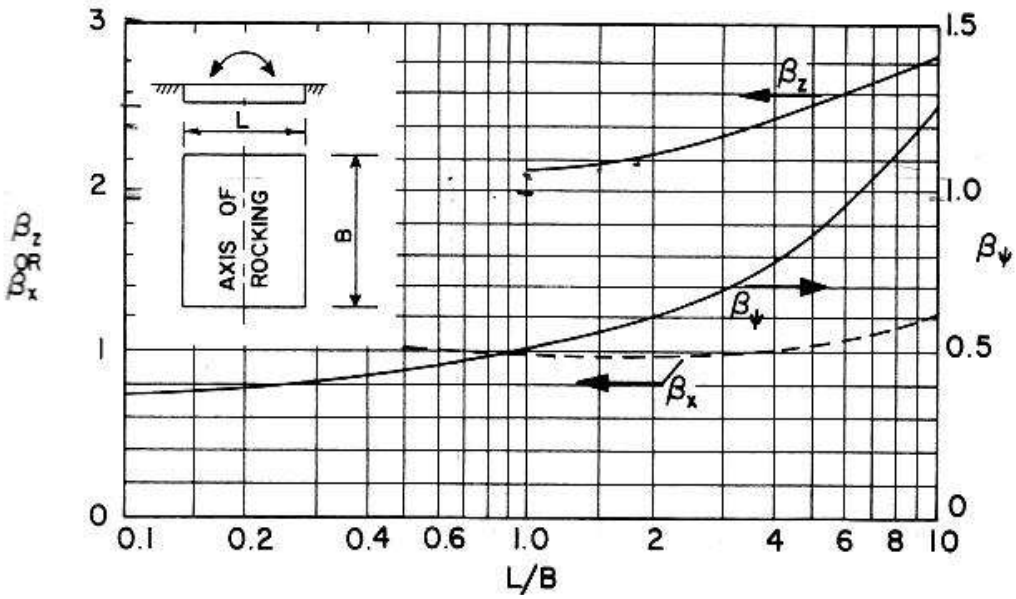


Fig. (2) – Coefficient β_z, β_x and β_ψ for rectangular footing (after Whitman, 1966).

To calculate the geometric damping ratio for vertical direction D_{g_z}, the damping ratio embedment factor for vertical direction $\alpha_{(z)}$ and mass ratio for vertical direction (B_z) have to be specified, while internal damping ratio (D_i) equals approximately (0.05) as follows (Das, 1983):

$$\alpha_{(z)} = \frac{1 + 1.9(1-\upsilon)\dfrac{h_0}{r_{oz}}}{\sqrt{\eta_z}} \tag{4}$$

$$B_z = \frac{(1-\upsilon)W_t}{4\gamma\ (r_{oz})^3} \tag{5}$$

$$D_{g_z} = \frac{0.425\alpha(z)}{\sqrt{B_z}} \tag{6}$$

The summation of geometric and internal damping gives total damping which contributes to the calculation of resonance frequency if resonance is possible or not depending on the term ($2D^2$) in the equation of resonance frequency eq. (7b) after calculating natural frequency eq. (7a) as given below (Das, 1983):

$$f_{nz} = \frac{1}{2\pi}\sqrt{\frac{\kappa_z}{m_t}} \tag{7a}$$

$$f_{m_z} = \frac{f_n}{\sqrt{1-2D^2}} \tag{7b}$$

After the resonance conditions are defined, the magnification factor (M_Z) should be calculated. The magnification factor is defined as the ratio of a steady-state displacement response caused by dynamic force (A_{max}) to the displacement caused by an equivalent static force of amplitude equals to the amplitude of the dynamic force (A_s) **Fig. (1)**:

$$M_Z = A_{max} / A_s \tag{8a}$$

$$\text{or } M_z = \frac{1}{\left(\left(1-\left(\frac{\omega}{\omega_n}\right)^2\right)^2 + \left(2D\frac{\omega}{\omega_n}\right)^2\right)^{1/2}} \tag{8b}$$

where (ω/ω_n) is the ratio of operating frequency to natural circular frequency (rps) in vertical direction which is calculated from:

$$\omega = 2\pi f \tag{9a}$$

$$\omega_n = 2\pi f_n \tag{9b}$$

in which f and f_n are the operating frequency of machine and natural frequency, respectively.

After all that, the displacement which occurs as a result of vibration is calculated depending on the vibration force obtained from the force diagrams that are usually supplied by the manufacturer of the machine as follows:

$$Z = \frac{M_z F_o}{\kappa_z} \tag{10}$$

where: F_0 is the amplitude of excitation force.

Then the transmissibility factor (T_r), which is defined as, "the ratio of the magnitude of the force transmitted to that of the impressed force", is calculated as follows (see **Fig. (3) and (4)**):

$$T_r = \frac{\sqrt{1+\left(2Dr^2\right)}}{\sqrt{\left(1-r^2\right)+\left(2Dr^2\right)}} \tag{11}$$

In the final step for design criteria, the transmitted force P_v is calculated as follows:

$$P_v = T_r F_o \tag{12}$$

These calculations will be carried out using the computer program MATHCAD. The results obtained by this procedure have to be compared with the design limits as shown in **Fig. (3)** in order to get the appropriate decision of design.

The permissible amplitudes of a machine foundation is governed by the relative importance of the machine and the sensitivity of neighboring structures to vibration. These limits are summarized in **Table (3)**.

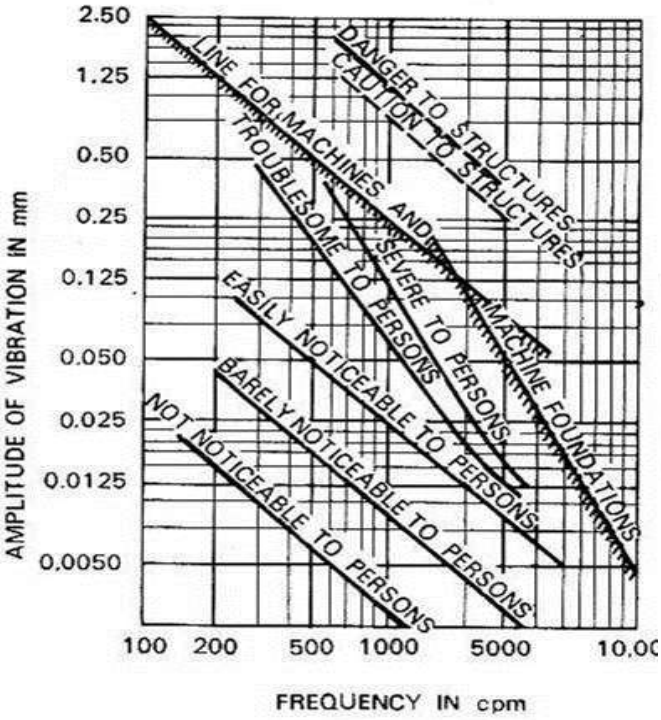


Fig. (3) – General limits of vibration amplitude (Richart, 1960).

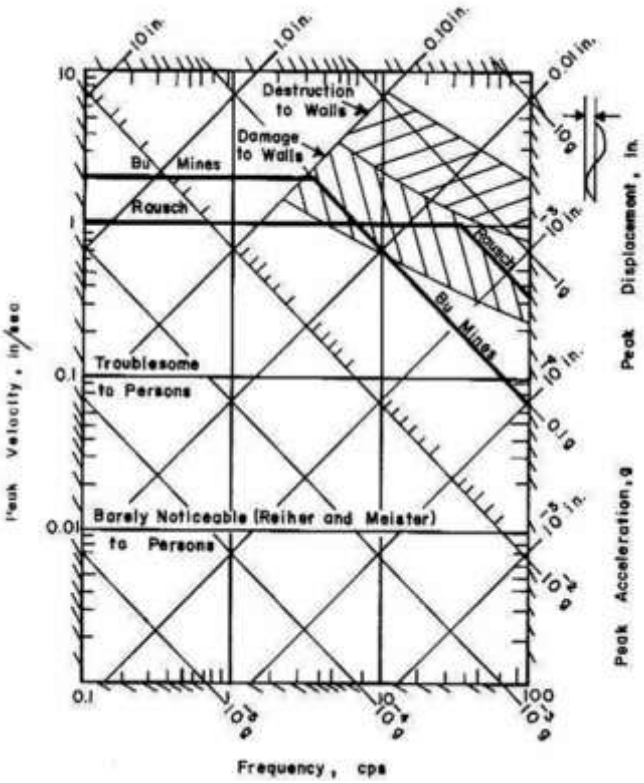


Fig. (4) – Response spectra for allowable vibration (Richart et al., 1970).

Table (3) – The permissible amplitudes of the machine foundation (Srinivasulu and Viadyanathan, 1976).

Type	**Permissible amplitudes** (m)
Low-speed machinery (500 rpm)	0.0002 to 0.00025
Hammer foundations	0.001 to 0.0012
High -speed machinery: a. (3000 rpm) b. (1500 rpm)	 0.00002 to 0.00003 0.00004 to 0.00006

THE COMPUTER PROGRAM MATHCAD

In order to apply the empirical method, the design equations need to be used more than one time for a given data. So to solve these equations with a little effort, time, and high accuracy, it is preferred to use assistant program. The computer software **MATHCAD** is used for this purpose.

MATHCAD program is a professional quality tool being increasingly used by many of scientists and engineers in the visualization and application of mathematics (Desrues, 1997). It is

the industry standard calculation software for technical professionals, educators, and college students. By using **MATHCAD** in calculating, the results become easy to understand.

DESIGN CHARTS FOR MACHINE FOUNDATIONS

To use of the solution presented in equations of machine foundation by the empirical method, design charts are prepared to be a guide for the designer engineer. The selected values used in these charts were limited based on the conditions considered in **Table (4)** as well as the limitations considered in the limitations of machine foundation. The design charts are selected based on three displacements which are acceptable for design of machine foundations as considered in **Fig. (3)** (Bowles, 1988).

Table (4): The parameters of the empirical method.

Parameters	Basic values	Range of values	Units
$W_{mach.}$	1444.905	60-620	kN
f	585	50-1000	rpm
γ	18.33	18-22	kN/m^3
G	96365	25000-190000	kN/m^2
ν	0.35	0.3-0.45	_
D_i	0.05	0.05-0.15	_
L_f	8.39	2-20	m
B_f	4.80	2-20	m
h	1.52	0.6-2.2	m
$W_{fou.}$	1443.95	$(3\text{-}5)W_{mach.}$	kN

For these displacements, the analysis is carried out using the computer program MATHCAD and the results are presented in the form of a relationship between ($G / \gamma L_f$) (y-axis) and frequency (rpm) (x-axis), for different ratios of the weight of the foundation to the weight of the machine ((W_f / W_m) W_f = weight of foundation, W_m = weight of machine) ranging between (3-5).

The selected displacement values ranged between (2.5×10^{-6}) m to a maximum value of (125×10^{-6}) m.

The charts are used to design the dimensions of the footing by the empirical method depending on the weight of the machine, the operating frequency of the machine and the properties of the soil including (shear modulus, Poisson's ratio and unit weight of the soil). In this paper we will take the effects of the minimum displacement = (2.5×10^{-6} m) on the design charts.

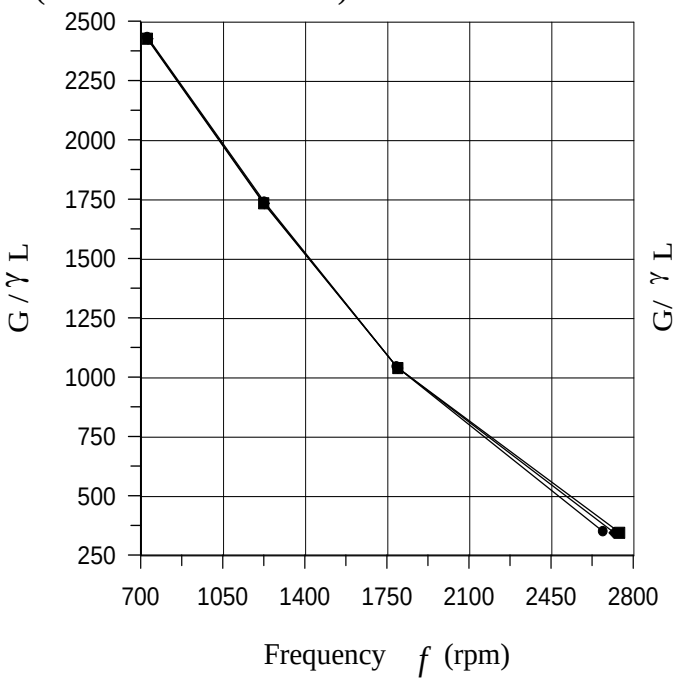
MINIMUM DISPLACEMENT = 2.5×10^{-6} m:

This displacement is considered for general limits of vibration which is not noticeable to persons as shown in **Fig. (3)**. **Fig. (5)** is drawn for the foundation dimensions ratio $L_f/B_f = 1$, Poisson's ratio, $\nu = 0.35$, and different soil unit weights, $\gamma = 18$, 20 and 22 kN /m$^{3.}$

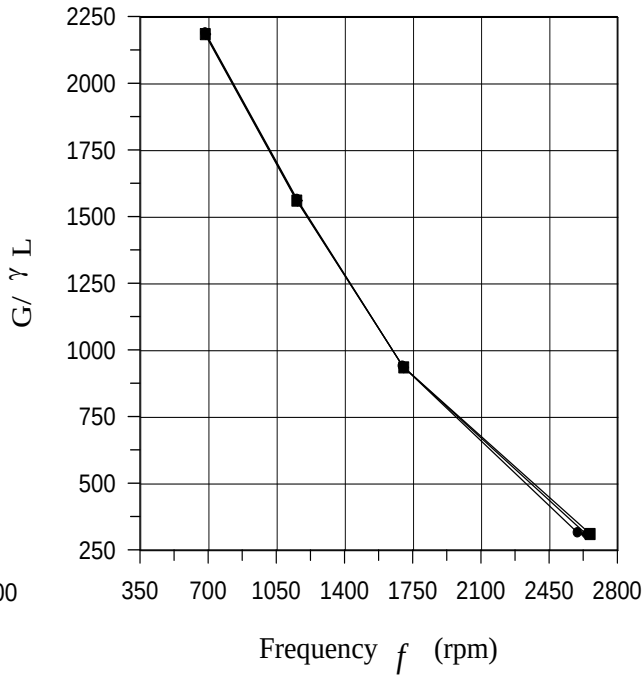
From these figures, it is apparent that the frequency is inversely proportional to the values of ($G/\gamma L_f$). The curves of these relationships for different values of (W_f / W_m) coincide with each other especially at frequency level (500-1750 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

The values of the shear modulus (G) used in these figures ranged between

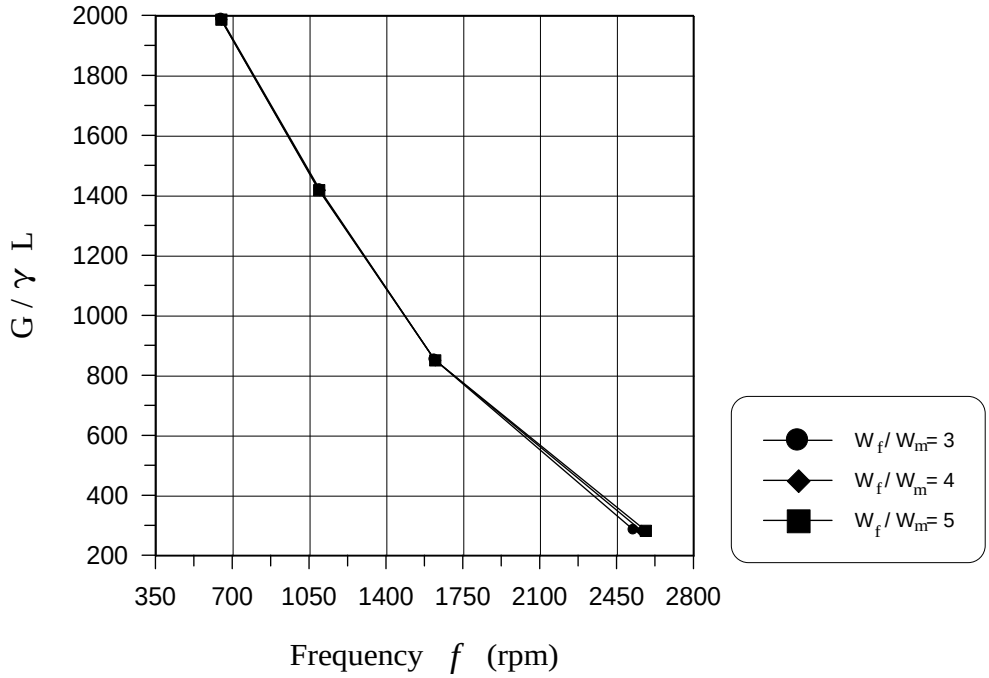
(25×10^3 and 175×10^3) kN/m^2.



a) $\gamma = 18$ kN/m^3



b) $\gamma = 20$ kN/m^3



c) $\gamma = 22$ kN/m^3

Fig. (6) is drawn for the foundati $B_f = 2$, Poisson's ratio, $\nu = 0.35$, and different soil unit weights, $\gamma = 18$, 20 ar

Fig. (5) ─ Design charts for machine foundations (L/B = 1, ν = 0.35) and displacement = 2.5 x 10^{-6} m.

From these figures, it is apparent that the frequency is also inversely proportional to the values of ($G / \gamma L_f$). The curves of these relationships for different values of (W_f / W_m) coincide with each other especially at frequency level (200-900 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

The values of the shear modulus (G) used in these figures ranged between (25×10^3 and 125×10^3) kN/m^2 because in the case of shear modulus equals to (175×10^3) kN/m^2, the resulted displacements were out of the limit of (2.5×10^{-6}) m.

Fig. (7) is drawn for the foundation dimensions ratio $L_f / B_f = 3$, Poisson's ratio, $\nu = 0.35$, and different soil unit weights, $\gamma = 18$, 20 and 22 $kN/m^{3.}$

From these figures, it is apparent that the frequency is inversely proportional to the values of ($G / \gamma L_f$). The curves of these relationships for different values of (W_f / W_m) coincide with each other especially at frequency level (200-1750 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

The values of the shear modulus (G) used in these figures ranged between (25×10^3 and 175×10^3) kN/m^2 except for $\gamma = 18$, the shear modulus (G) ranged between (25×10^3 and 125×10^3) kN/m^2.

Fig. (8) is drawn for the foundation dimensions ratio $L_f / B_f = 1$, Poisson's ratio, $\nu = 0.4$, and different soil unit weights, $\gamma = 18$, 20 and 22 $kN/m^{3.}$

As in the previous figures, it is apparent that the frequency is inversely proportional to the values of ($G / \gamma L_f$). The curves of these relationships for different values of (W_f / W_m) coincide with each other especially at frequency level (400-2000 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

The values of the shear modulus (G) used in these figures ranged between (25×10^3 and 125×10^3) kN/m^2 except for $\gamma = 18$ kN/m^3, the shear modulus (G) ranged between (25×10^3 and 175×10^3) kN/m^2.

Fig. (9) is drawn for the foundation dimensions ratio $L_f / B_f = 2$, Poisson's ratio, $\nu = 0.4$, and different soil unit weights, $\gamma = 18$, 20 and 22 $kN/m^{3.}$

The same relationship between the frequency and the values of ($G / \gamma L_f$). The curves of these relationships for different values of (W_f / W_m) coincide with each other especially at frequency level (400-1000 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

The values of the shear modulus (G) used in these figures ranged (25×10^3 and 125×10^3) kN/m^2.

Fig. (10) is drawn for the foundation dimensions ratio $L_f / B_f = 3$, Poisson's ratio, $\nu = 0.4$, and different soil unit weights, $\gamma = 18$, 20 and 22 $kN/m^{3.}$

From these figures, it is apparent that the frequency is inversely proportional to the values of ($G / \gamma L_f$). The curves of these relationships for different values of (W_f / W_m) coincide with each other especially at frequency level (500-700 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

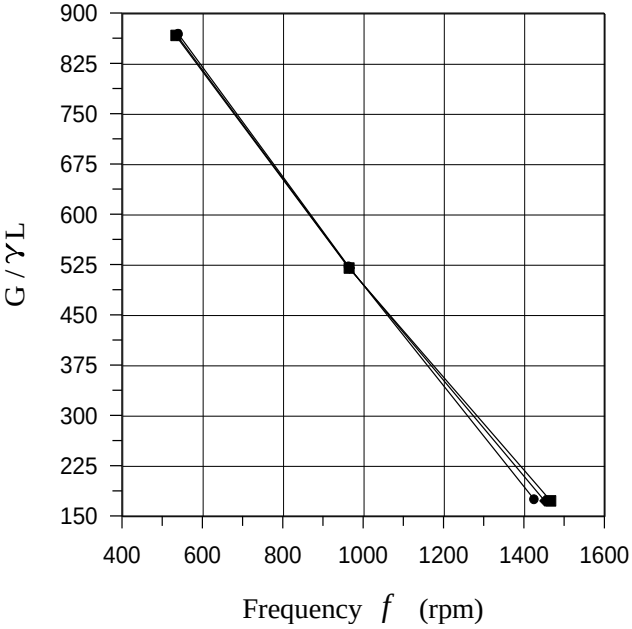
The values of the shear modulus (G) used in these figures are (25×10^3 and 75×10^3) kN/m^2.

Fig. (11) is drawn for the foundation dimensions ratio $L_f / B_f = 1$, Poisson's ratio, $\nu = 0.45$, and different soil unit weights, $\gamma = 18$, 20 and 22 $kN/m^{3.}$

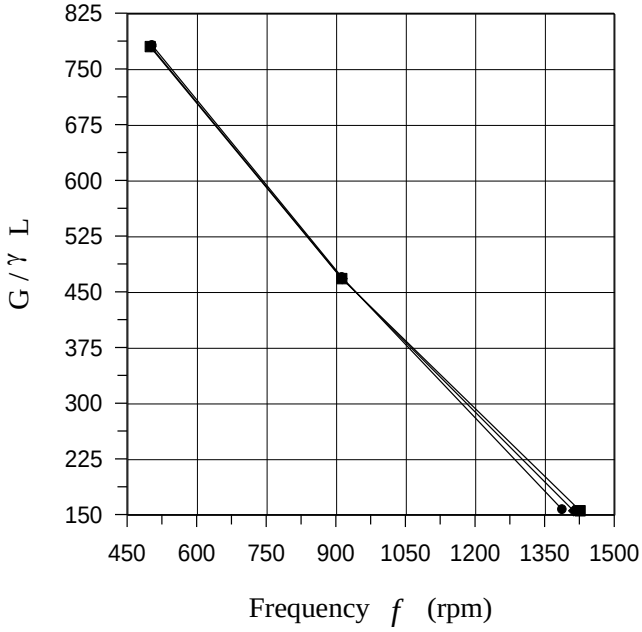
From these figures, it is apparent that the frequency is inversely proportional to the values of ($G / \gamma L_f$). The curves of these relationships for different values of (W_f / W_m) coincide with each

other especially at frequency level (250-2000 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

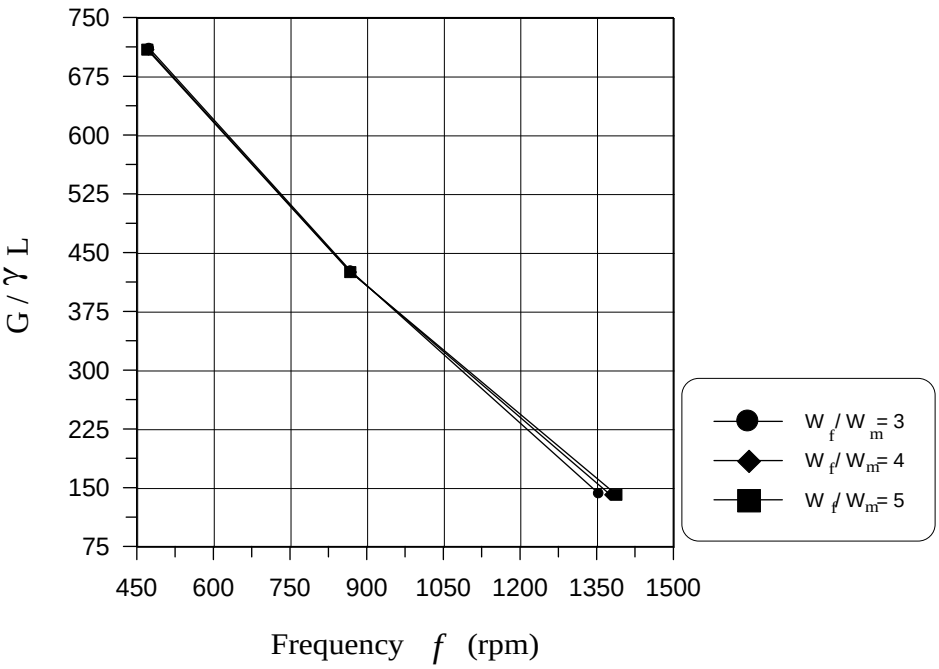
The values of the shear modulus (G) used in these figures ranged between (25 x 10^3 and 175 x 10^3) kN/m^2.



a) γ = 18 kN/m^3



b) γ = 20 kN/m^3



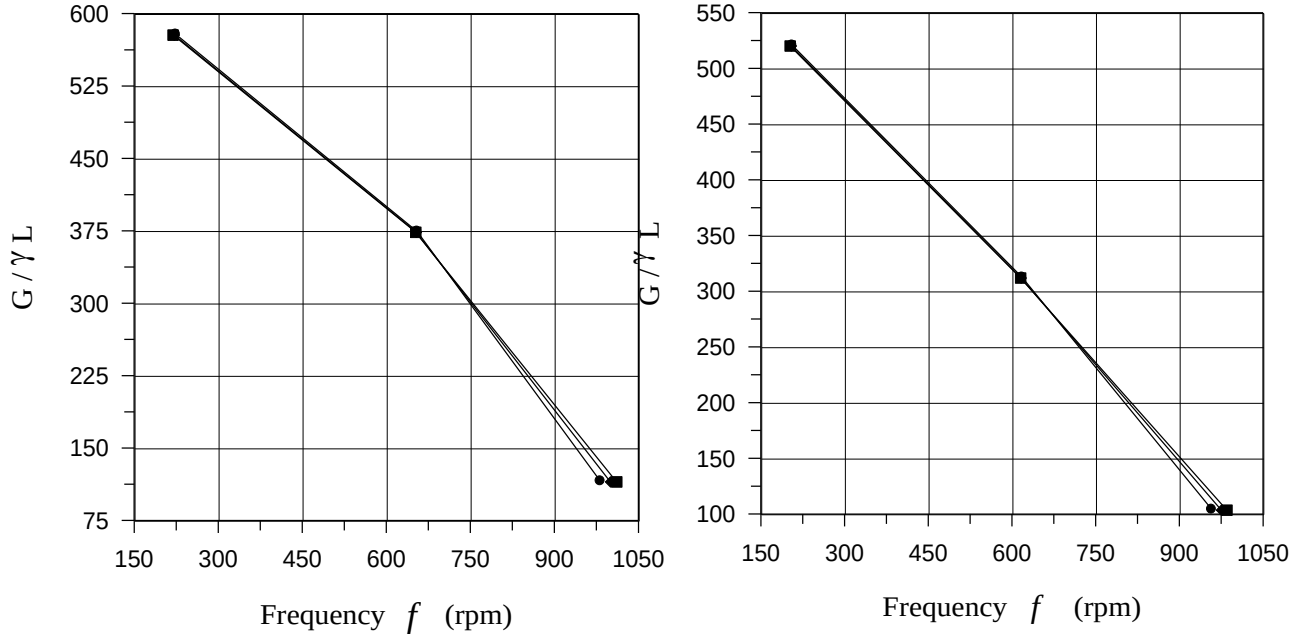
c) γ = 22 kN/m^3

Fig. (6) – Design charts for machine foundations (L/B = 2, ν = 0.35) and displacement = 2.5 x 10^{-6} m.

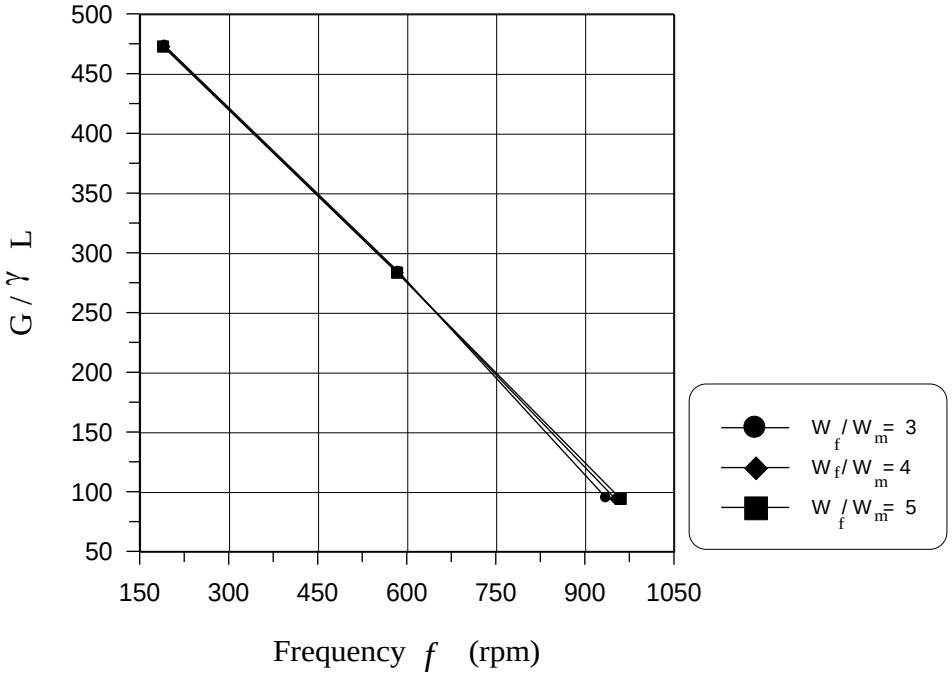
Fig. (12) is drawn for the foundation dimensions ratio $L_f / B_f = 2$, Poisson's ratio, $\nu = 0.45$, and different soil unit weights, $\gamma = 18$, 20 and 22 kN /m$^{3.}$

From these figures, it is apparent that the frequency is inversely proportional to the values of ($G / \gamma L_f$). The curves of these relationships for different values of (W_f / W_m) also coincide with each other especially at frequency level (200-1000 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

The values of the shear modulus (G) used in these figures ranged between (25 x 10^3 and 75 x 10^3) kN/m^2.



a) $\gamma = 18$ kN/m^3 **b) $\gamma = 20$ kN/m^3**



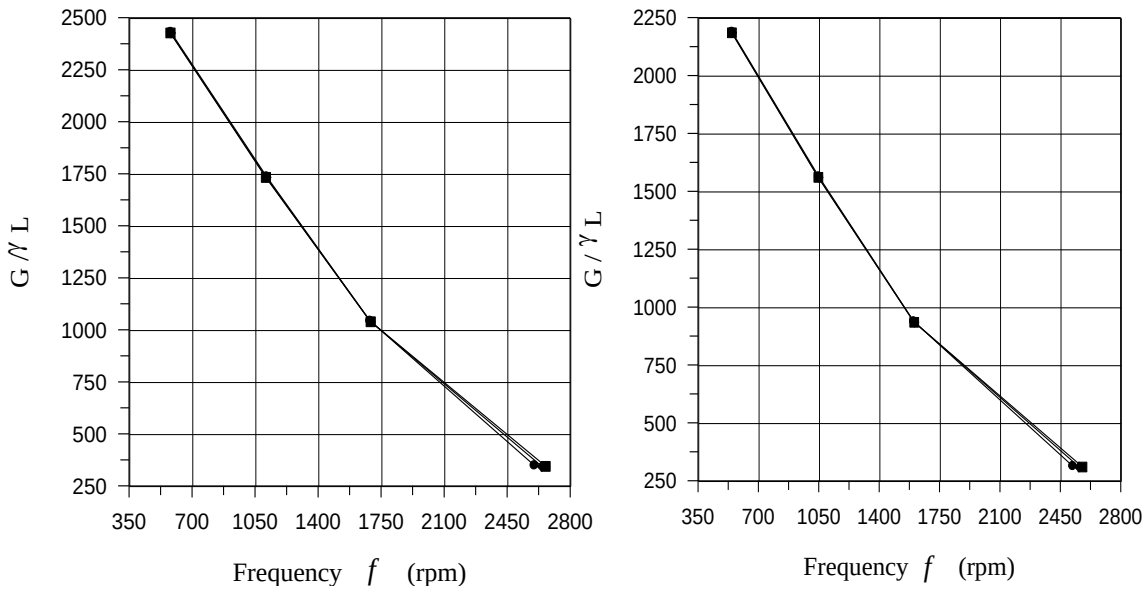
c) $\gamma = 22$ kN/m^3

Fig. (7) ─ Design charts for machine foundations (L/B = 3, ν = 0.35) and displacement = 2.5 x 10^{-6} m.

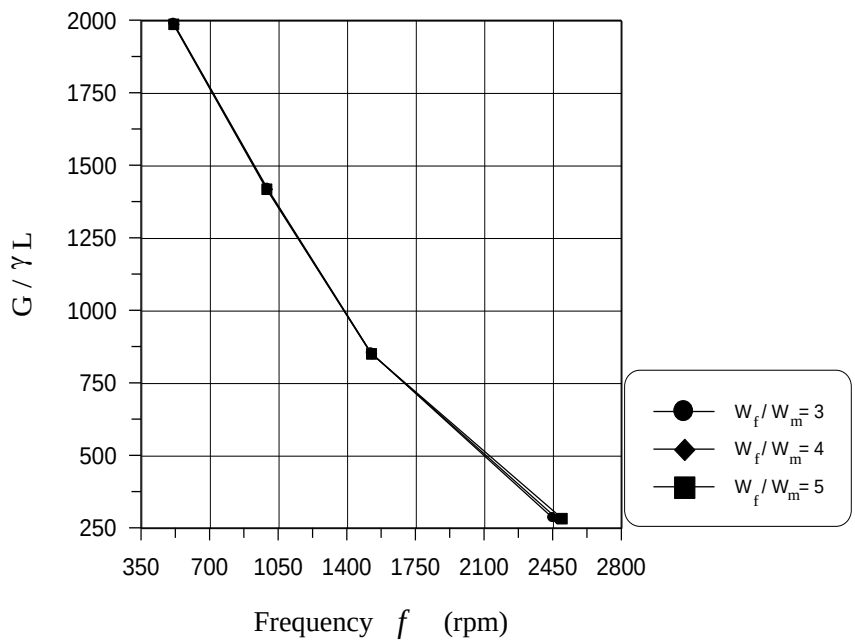
Fig. (13) is drawn for the foundation dimensions ratio $L_f / B_f = 3$, Poisson's ratio, $\nu = 0.45$, and different soil unit weights, $\gamma = 18$, 20 and 22 kN /m$^{3.}$

From these figures, it is apparent that the frequency is inversely proportional to the values of ($G / \gamma L_f$). The curves of these relationships for different values of (W_f / W_m) coincide with each other especially at frequency level (500-650 rpm). After this limit of frequency, the effect of the weight ratio can be pronounced.

The values of the shear modulus (G) used in these figures are (25 x 10^3 and 75 x 10^3) kN/m^2.

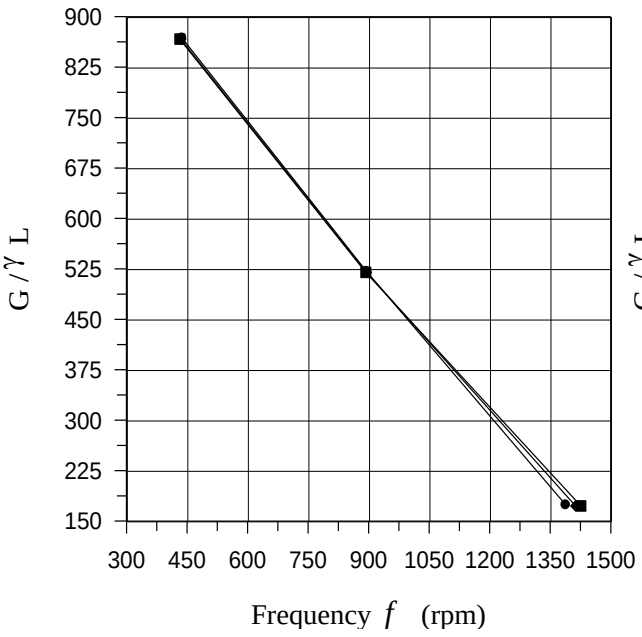


a) γ = 18 kN/m^3 **b) γ = 20 kN/m^3**

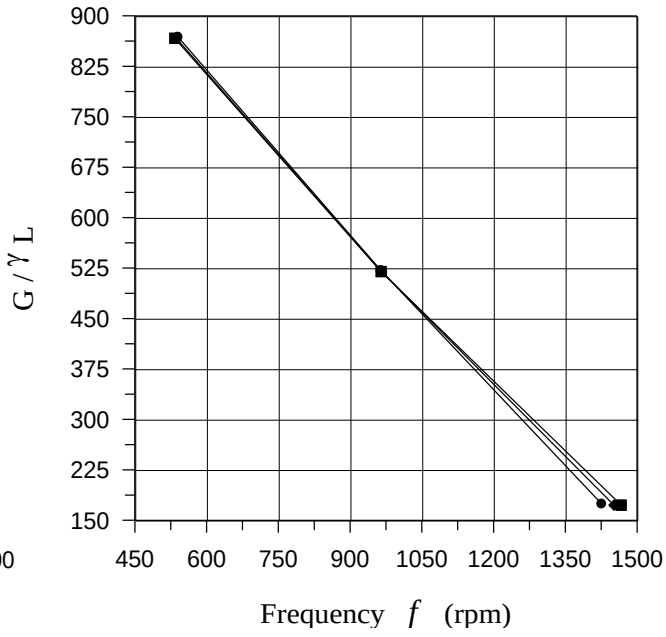


c) γ = 22 kN/m^3

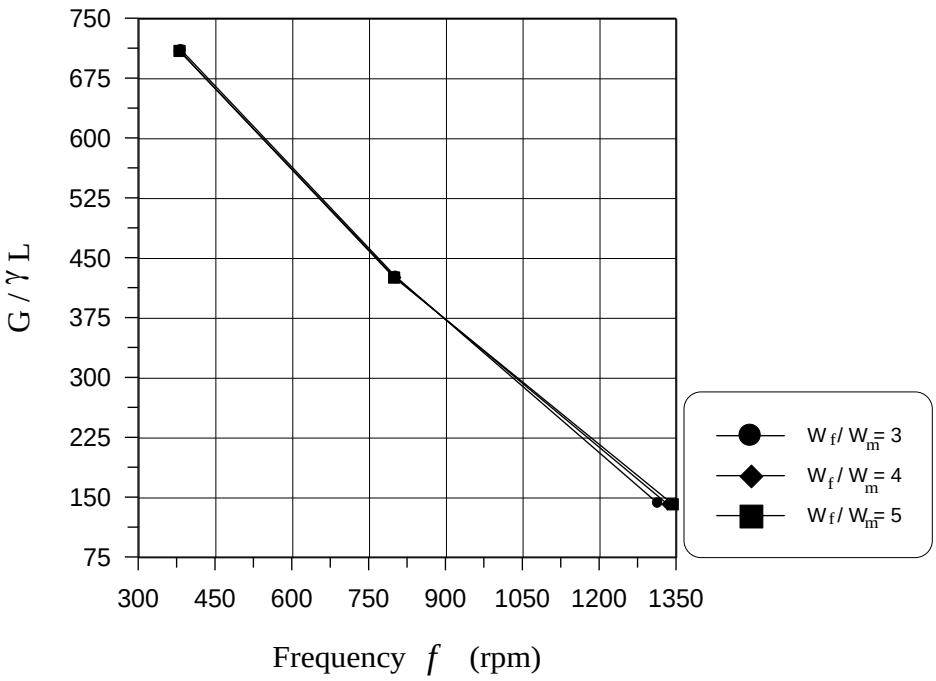
Fig. (8) ─ Design charts for machine foundations (L/B = 1, ν = 0.4) and displacement = 2.5 x 10^{-6} m.



a) γ = 18 kN/m^3

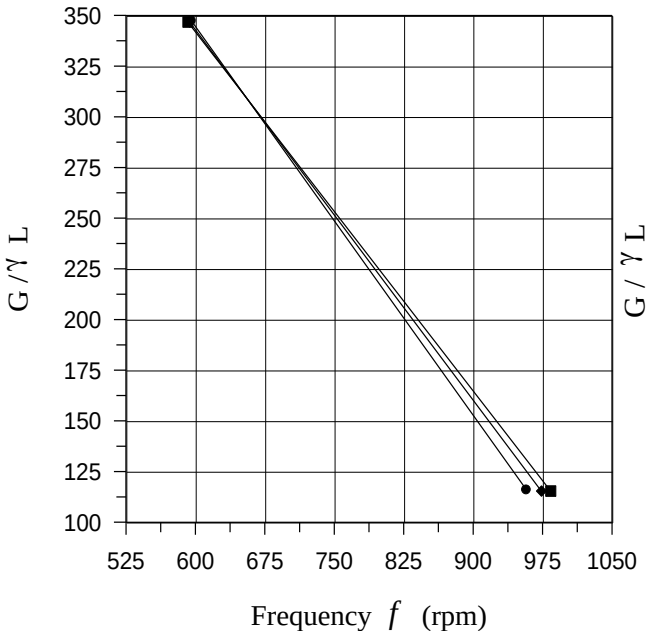


b) γ = 20 kN/m^3

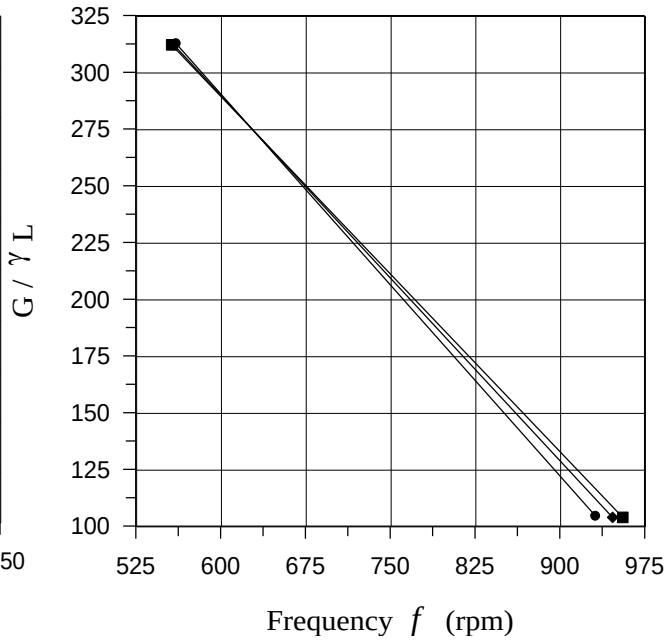


c) γ = 22 kN/m^3

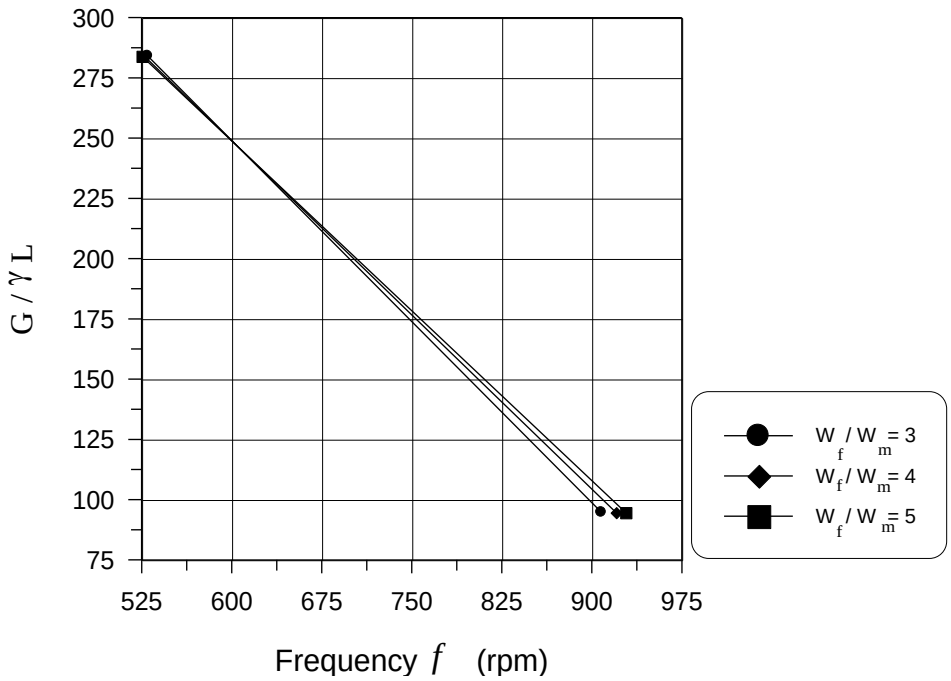
Fig. (9) ─ Design charts for machine foundations (L/B = 2, ν = 0.4) and displacement = 2.5 x 10^{-6} m.



a) γ = 18 kN/m^3

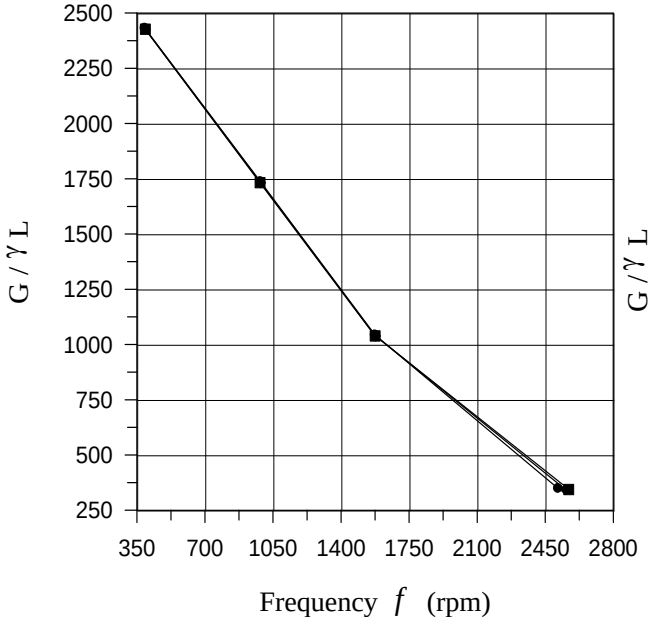


b) γ = 20 kN/m^3

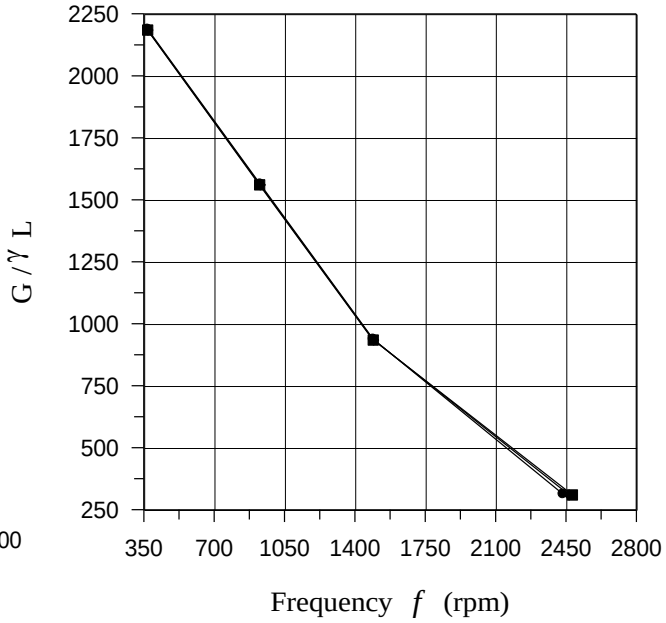


c) γ = 22 kN/m^3

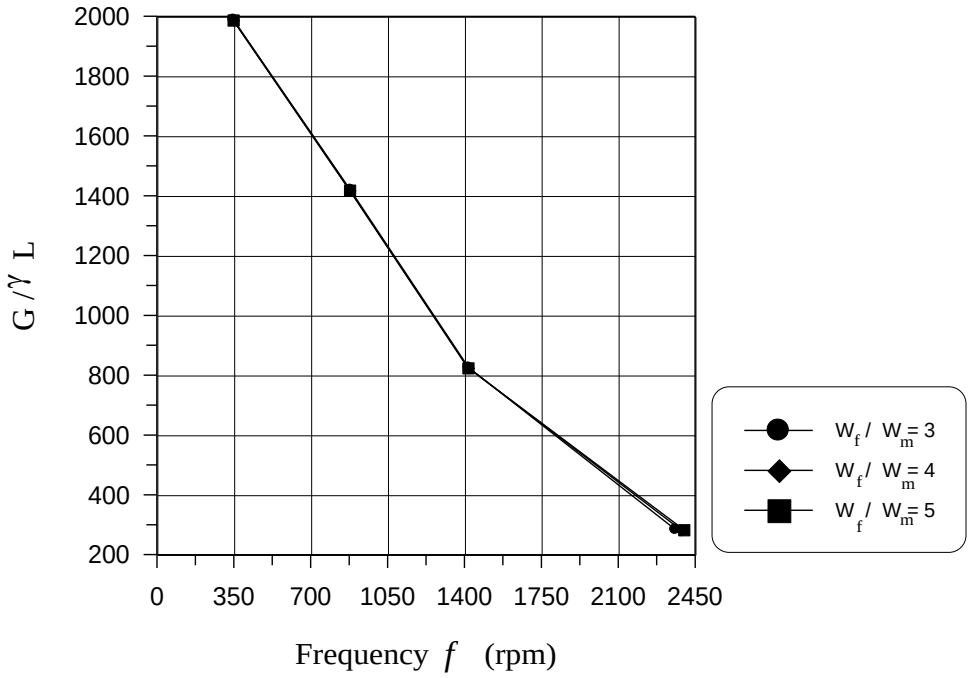
Fig. (10) ─ Design charts for machine foundations (L/B = 3, ν = 0.4) and displacement = 2.5 x 10^{-6} m.



a) γ = 18 kN/m^3

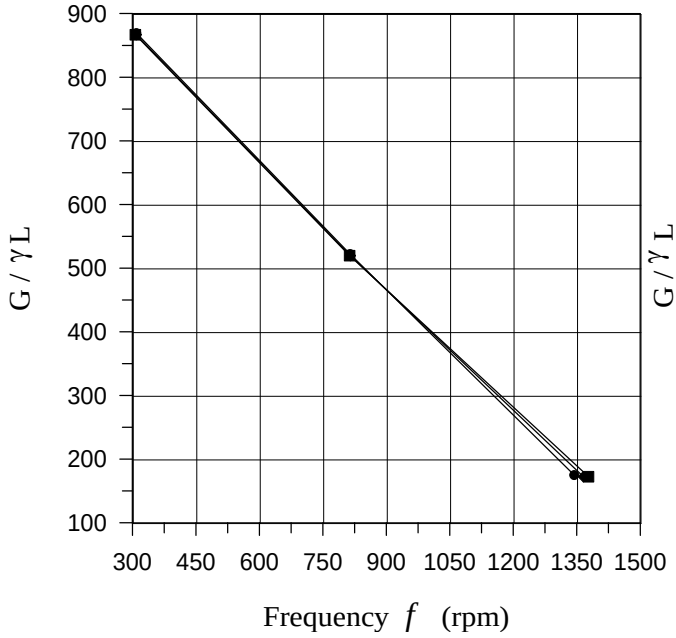


b) γ = 20 kN/m^3

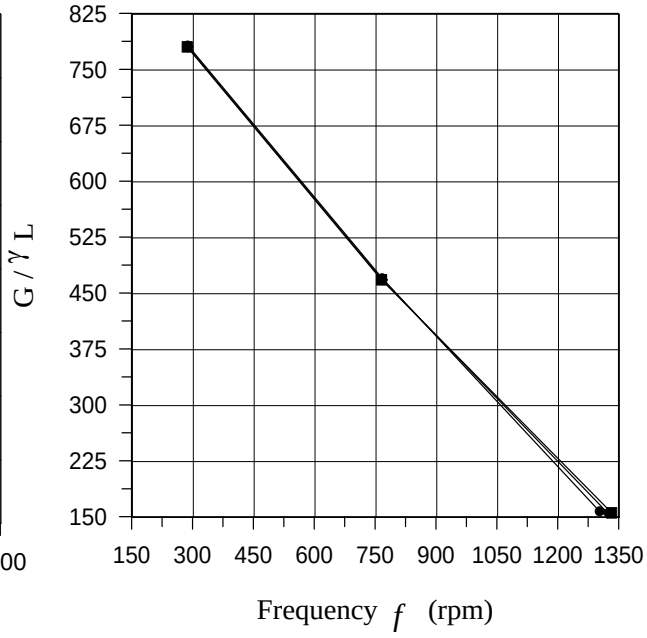


c) γ = 22 kN/m^3

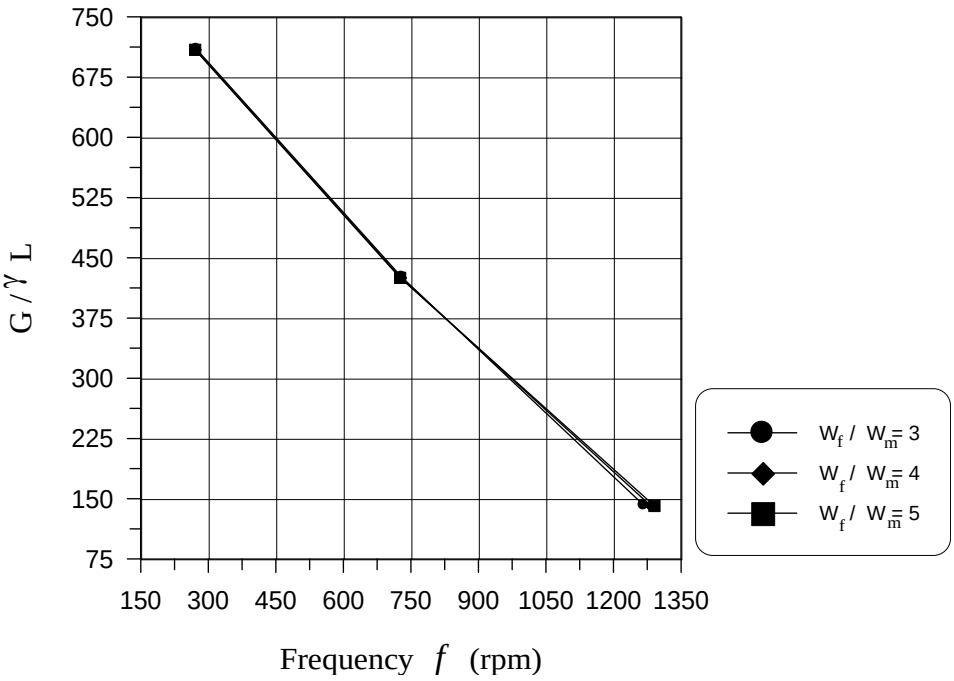
Fig. (11) ─ Design charts for machine foundations (L/B = 1, ν = 0.45) and displacement = 2.5 x 10^{-6} m.



a) γ = 18 kN/m^3

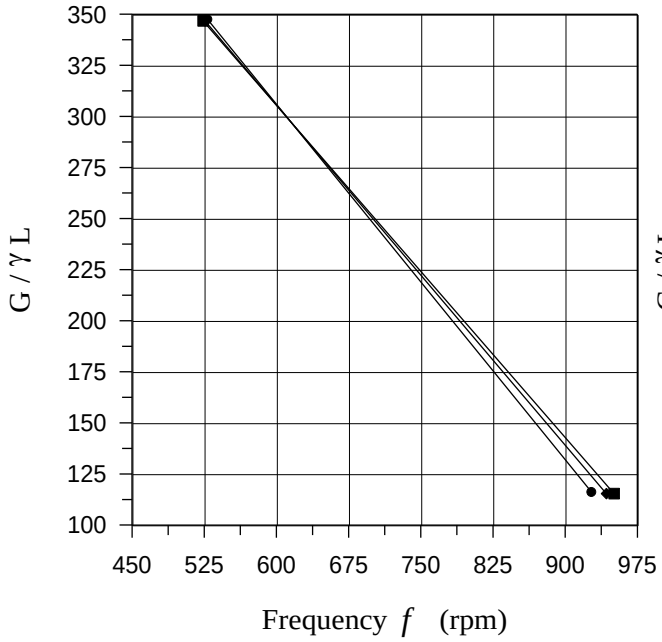


b) γ = 20 kN/m^3

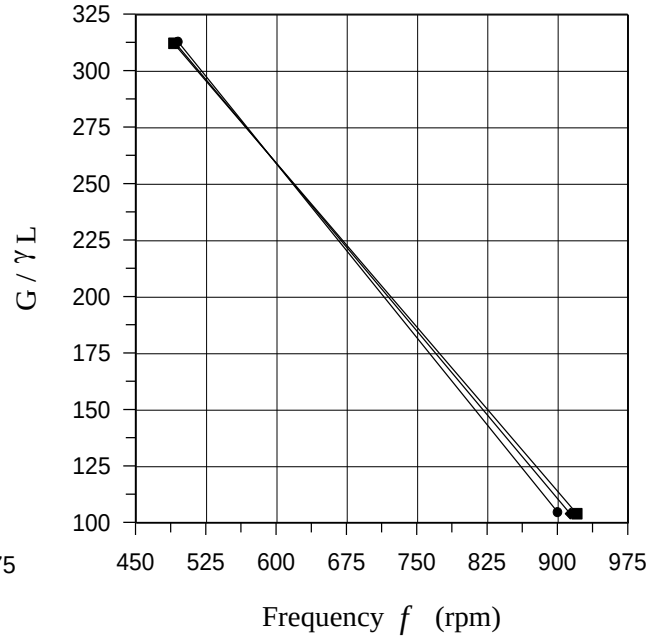


c) γ = 22 kN/m^3

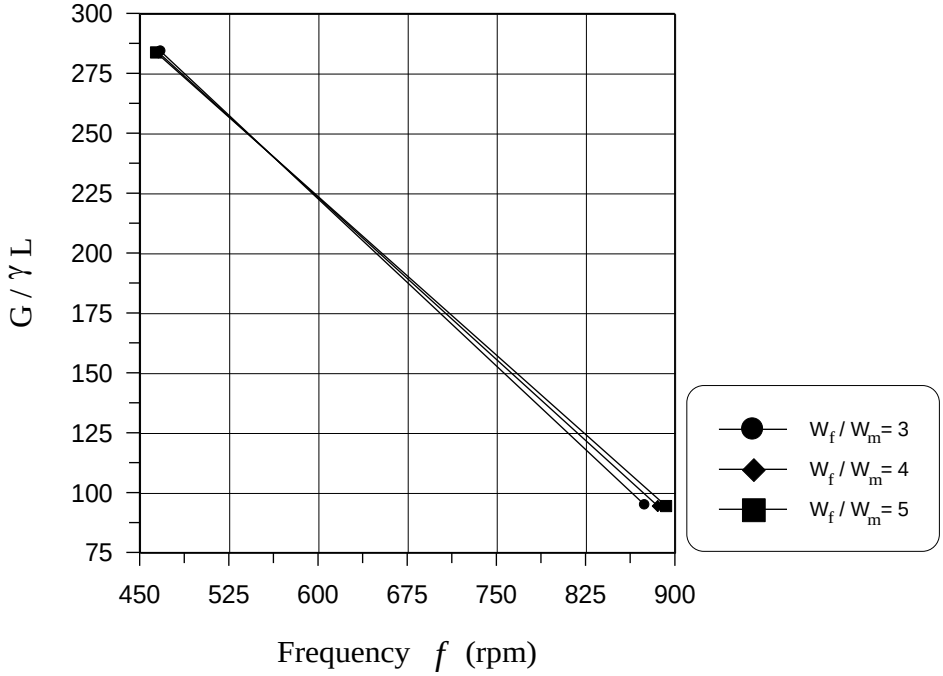
Fig. (12) – Design charts for machine foundations (L/B = 2, ν = 0.45) and displacement = 2.5 x 10^{-6} m.



a) γ = 18 kN/m^3



b) γ = 20 kN/m^3



c) γ = 22 kN/m^3

Fig. (13) ─ Design charts for machine foundations (L/B = 3, ν = 0.45) and displacement = 2.5 x 10^{-6} m.

CONCLUSIONS

It was found that the most important variable affecting the problem of machine foundations is the shear modulus of the soil. Considering the shear modulus as state variable, it was found that by the empirical method, the maximum displacement decreases when the shear modulus increases as the type of soil is sand; and the maximum displacement is smaller than the case when the type of soil is clay. For the cone model method, the maximum displacement decreases when the shear modulus increases when the shear modulus is less than 200 kN/m^2 for the range of soils analyzed in this study, while when the shear modulus is more than 200 kN/m^2, the maximum displacement increases with the increase of the shear modulus.

The maximum displacement decreases with the increase of machine operating frequency, soil unit weight, shear modulus, Poisson's ratio and internal damping.

REFERENCES

- Arya, S. C., O'Neill, M. W. and Pincus, G., (1979), ***"Design of Structures and Foundations for Vibrating Machine"***, Gulf Publishing Company Book Division, Huston, London, Tokyo.

- Bowles, J. E., (1988), ***"Foundation analysis and design"***, 4th ed., McGraw-Hill, New York, 1004 pp.

- Das, B. M., (1983). ***" Fundamentals of Soil Dynamics",*** Elsevier Science Publishing Co., Inc.

- Desrues, K. P., (1997), ***"Exploration of Mathcad",*** Addison-Wesley Publishers Lid., 227 pp.

- Gazetas, G., (1983), ***"Analysis of Machine Foundation Vibrations: state of the art"***, Journal of Soil Dynamics and Earthquake Engineering, chapter- 2, pp. 2-42.

- Gazetas, G. and Rosset, J.M., (1979), ***"Vertical Vibration of Machine Foundations"***, Journal of Geotechnical Engineering, ASCE, 105(12), pp. 1435-1454.

- Harr, M. H., (1966), ***"Foundations of Theoretical Soil Mechanics"***, McGraw-Hill, New York.

- Reissner, E., (1936), ***"Station are Axialsymmetricle Ddruch Eine Elastischen HalbRraues",*** Ingenieur Archiv, Vol. 7, part-6, pp. 381-396, (as cited by Pradhan et al., 2003).

- Richart, F. E. Jr., Hall, J. R. Jr. and Woods, R. D., (1970), ***"Vibrations of Soils and Foundations"***, Prentice-Hall, Inc. Englewood Cliffs, New Jersey.

- Srinivasulu, P. and Vaidyanathan, C. V., (1976). ***"Machine Foundations",*** Structural

Engineering Research Center, Roorkee-Madras.

- Whitman, R. V., (1966), ***"Analysis of Foundation Vibrations",*** Vibration in Civil Engineering, B. O. Skip (ed.), Butterworths, London, pp. 159-179.